

# On the entropy numbers of the mixed smoothness function classes

V. Temlyakov \*

## Abstract

Behavior of the entropy numbers of classes of multivariate functions with mixed smoothness is studied here. This problem has a long history and some fundamental problems in the area are still open. The main goal of this paper is to develop a new method of proving the upper bounds for the entropy numbers. This method is based on recent developments of nonlinear approximation, in particular, on greedy approximation. This method consists of the following two steps strategy. At the first step we obtain bounds of the best  $m$ -term approximations with respect to a dictionary. At the second step we use general inequalities relating the entropy numbers to the best  $m$ -term approximations. For the lower bounds we use the volume estimates method, which is a well known powerful method for proving the lower bounds for the entropy numbers. It was used in a number of previous papers.

## 1 Introduction

Behavior of the entropy numbers of classes of multivariate functions with mixed smoothness is studied here. This problem has a long history and some fundamental problems in the area are still open. The main goal of this paper is to develop a new method of proving the upper bounds for the entropy numbers. This method is based on recent developments of nonlinear approximation, in particular, on greedy approximation. This method consists

---

\*University of South Carolina and Steklov Institute of Mathematics.

of the following two steps strategy. At the first step we obtain bounds of the best  $m$ -term approximations with respect to a dictionary. At the second step we use general inequalities relating the entropy numbers to the best  $m$ -term approximations. For the lower bounds we use the volume estimates method, which is a well known powerful method for proving the lower bounds for the entropy numbers. It was used in a number of previous papers. Taking into account the fact that there are fundamental open problems in the area, we give a detailed discussion of known results and of open problems. We also provide some comments on the techniques, which were used to obtain known results. Then we formulate our new results and compare them to the known results.

Let  $X$  be a Banach space and let  $B_X$  denote the unit ball of  $X$  with the center at 0. Denote by  $B_X(y, r)$  a ball with center  $y$  and radius  $r$ :  $\{x \in X : \|x - y\| \leq r\}$ . For a compact set  $A$  and a positive number  $\varepsilon$  we define the covering number  $N_\varepsilon(A)$  as follows

$$N_\varepsilon(A) := N_\varepsilon(A, X) := \min\{n : \exists y^1, \dots, y^n : A \subseteq \cup_{j=1}^n B_X(y^j, \varepsilon)\}.$$

It is convenient to consider along with the entropy  $H_\varepsilon(A, X) := \log_2 N_\varepsilon(A, X)$  the entropy numbers  $\varepsilon_k(A, X)$ :

$$\varepsilon_k(A, X) := \inf\{\varepsilon : \exists y^1, \dots, y^{2^k} \in X : A \subseteq \cup_{j=1}^{2^k} B_X(y^j, \varepsilon)\}.$$

Let  $F_r(x, \alpha)$  be the univariate Bernoulli kernels

$$F_r(x, \alpha) := 1 + 2 \sum_{k=1}^{\infty} k^{-r} \cos(kx - \alpha\pi/2).$$

For  $\mathbf{x} = (x_1, \dots, x_d)$  and  $\alpha = (\alpha_1, \dots, \alpha_d)$  we define

$$F_r(\mathbf{x}, \alpha) := \prod_{i=1}^d F_r(x_i, \alpha_i)$$

and

$$\mathbf{W}_{q,\alpha}^r := \{f : f = F_r(\cdot, \alpha) * \varphi, \quad \|\varphi\|_q \leq 1\}$$

where  $*$  means convolution. In the univariate case we use the notation  $W_{q,\alpha}^r$ .

It is well known that in the univariate case

$$\varepsilon_k(W_{q,\alpha}^r, L_p) \asymp k^{-r} \tag{1.1}$$

holds for all  $1 \leq q, p \leq \infty$  and  $r > (1/q - 1/p)_+$ . We note that condition  $r > (1/q - 1/p)_+$  is a necessary and sufficient condition for compact embedding of  $W_{q,\alpha}^r$  into  $L_p$ . Thus (1.1) provides a complete description of the rate of  $\varepsilon_k(W_{q,\alpha}^r, L_p)$  in the univariate case. We point out that (1.1) shows that the rate of decay of  $\varepsilon_k(W_{q,\alpha}^r, L_p)$  depends only on  $r$  and does not depend on  $q$  and  $p$ . In this sense the strongest upper bound (for  $r > 1$ ) is  $\varepsilon_k(W_{1,\alpha}^r, L_\infty) \ll k^{-r}$  and the strongest lower bound is  $\varepsilon_k(W_{\infty,\alpha}^r, L_1) \gg k^{-r}$ .

There are different generalizations of classes  $W_{q,\alpha}^r$  to the case of multivariate functions. In this section we only discuss known results for classes  $\mathbf{W}_{q,\alpha}^r$  of functions with bounded mixed derivative. For further discussions see [30], Chapter 3 and [5].

The following two theorems are from [21] and [22].

**Theorem 1.1.** *For  $r > 1$  and  $1 < q, p < \infty$  one has*

$$\varepsilon_k(\mathbf{W}_{q,\alpha}^r, L_p) \ll k^{-r} (\log k)^{r(d-1)}.$$

**Theorem 1.2.** *For  $r > 0$  and  $1 \leq q < \infty$  one has*

$$\varepsilon_k(\mathbf{W}_{q,\alpha}^r, L_1) \gg k^{-r} (\log k)^{r(d-1)}.$$

The problem of estimating  $\varepsilon_k(\mathbf{W}_{q,\alpha}^r, L_p)$  has a long history. The first result on the right order of  $\varepsilon_k(\mathbf{W}_{2,\alpha}^r, L_2)$  was obtained by Smolyak [17]. Later (see [21], [22] and theorems above) it was established that

$$\varepsilon_k(\mathbf{W}_{q,\alpha}^r, L_p) \asymp k^{-r} (\log k)^{r(d-1)} \quad (1.2)$$

holds for all  $1 < q, p < \infty$ ,  $r > 1$ . The case  $1 < q = p < \infty$ ,  $r > 0$  was established by Dinh Dung [4]. Belinskii [2] extended (1.2) to the case  $r > (1/q - 1/p)_+$  when  $1 < q, p < \infty$ .

It is known in approximation theory (see [25]) that investigation of asymptotic characteristics of classes  $\mathbf{W}_{q,\alpha}^r$  in  $L_p$  becomes more difficult when  $q$  or  $p$  takes value 1 or  $\infty$  than when  $1 < q, p < \infty$ . It turns out to be the case for  $\varepsilon_k(\mathbf{W}_{q,\alpha}^r, L_p)$  too. It was discovered that in some of these extreme cases ( $q$  or  $p$  equals 1 or  $\infty$ ) relation (1.2) holds and in other cases it does not hold. We describe the picture in detail. It was proved in [22] that (1.2) holds for  $p = 1$ ,  $1 < q < \infty$ ,  $r > 0$ . It was also proved that (1.2) holds for  $p = 1$ ,  $q = \infty$  (see [2] for  $r > 1/2$  and [10] for  $r > 0$ ). Summarizing, we state that (1.2) holds for  $1 < q, p < \infty$  and  $p = 1$ ,  $1 < q \leq \infty$  for all  $d$  (with appropriate restrictions on  $r$ ). This easily implies that (1.2) also holds for  $q = \infty$ ,  $1 \leq p < \infty$ . For

all other pairs  $(q, p)$ , namely, for  $p = \infty$ ,  $1 \leq q \leq \infty$  and  $q = 1$ ,  $1 \leq p \leq \infty$  the rate of  $\varepsilon_k(\mathbf{W}_{q,\alpha}^r, L_p)$  is not known in the case  $d > 2$ . It is an outstanding open problem.

In the case  $d = 2$  this problem is essentially solved. We now cite the corresponding results. The first result on the right order of  $\varepsilon_k(\mathbf{W}_{q,\alpha}^r, L_p)$  in the case  $p = \infty$  was obtained by Kuelbs and Li [12] for  $q = 2$ ,  $r = 1$ . It was proved in [26] that

$$\varepsilon_k(\mathbf{W}_{q,\alpha}^r, L_\infty) \asymp k^{-r}(\log k)^{r+1/2} \quad (1.3)$$

holds for  $1 < q < \infty$ ,  $r > 1$ . We note that the upper bound in (1.3) was proved under condition  $r > 1$  and the lower bound in (1.3) was proved under condition  $r > 1/q$ . Belinskii [2] proved the upper bound in (1.3) for  $1 < q < \infty$  under condition  $r > \max(1/q, 1/2)$ . Relation (1.3) for  $q = \infty$  under assumption  $r > 1/2$  was proved in [28].

The case  $q = 1$ ,  $1 \leq p \leq \infty$  was settled by Kashin and Temlyakov [11]. The authors proved that

$$\varepsilon_k(\mathbf{W}_{1,\alpha}^r, L_p) \asymp k^{-r}(\log k)^{r+1/2} \quad (1.4)$$

holds for  $1 \leq p < \infty$ ,  $r > \max(1/2, 1 - 1/p)$  and

$$\varepsilon_k(\mathbf{W}_{1,0}^r, L_\infty) \asymp k^{-r}(\log k)^{r+1}, \quad r > 1. \quad (1.5)$$

Let us make an observation on the base of the above discussion. In the univariate case the entropy numbers  $\varepsilon_k(W_{q,\alpha}^r, L_p)$  have the same order of decay with respect to  $k$  for all pairs  $(q, p)$ ,  $1 \leq q, p \leq \infty$ . In the case  $d = 2$  we have three different orders of decay of  $\varepsilon_k(\mathbf{W}_{q,\alpha}^r, L_p)$  which depend on the pair  $(q, p)$ . For instance, in the case  $1 < q, p < \infty$  it is  $k^{-r}(\log k)^r$ , in the case  $q = 1$ ,  $1 < p < \infty$ , it is  $k^{-r}(\log k)^{r+1/2}$  and in the case  $q = 1$ ,  $p = \infty$  it is  $k^{-r}(\log k)^{r+1}$ .

We discussed above results on the right order of decay of the entropy numbers. Clearly, each order relation  $\asymp$  is a combination of the upper bound  $\ll$  and the matching lower bound  $\gg$ . We now briefly discuss methods that were used for proving upper and lower bounds. The upper bounds in Theorem 1.1 were proved by the standard method of reduction by discretization to estimates of the entropy numbers of finite-dimensional sets. Here results of [7], [14] or [16] are applied. It is clear from the above discussion that it was sufficient to prove the lower bound in (1.2) in the case  $p = 1$ . The proof

of this lower bound (see Theorem 1.2) is more difficult and is based on non-trivial estimates of the volumes of the sets of Fourier coefficients of bounded trigonometric polynomials. Theorem 2.4 (see below) plays a key role in this method.

An analogue of the upper bound in (1.3) for any  $d$  was obtained by Belinskii [2]: for  $q > 1$  and  $r > \max(1/q, 1/2)$  we have

$$\varepsilon_k(\mathbf{W}_{q,\alpha}^r, L_\infty) \ll k^{-r}(\log k)^{(d-1)r+1/2}. \quad (1.6)$$

That proof is based on Theorem 2.2 (see below).

Kuelbs and Li [12] discovered the fact that there is a tight relationship between small ball problem and the behavior of the entropy  $H_\varepsilon(\mathbf{W}_{2,\alpha}^1, L_\infty)$ . Based on results obtained by Livshits and Tsirelson [13], by Bass [1], and by Talagrand [18] for the small ball problem, they proved

$$\varepsilon_k(\mathbf{W}_{2,\alpha}^1, L_\infty) \asymp k^{-1}(\ln k)^{3/2}. \quad (1.7)$$

Proof of the most difficult part of (1.7) – the lower bound – is based on a special inequality, known now as the Small Ball Inequality, for the Haar polynomials proved by Talagrand [18] (see [27] for a simple proof).

We discussed above known results on the rate of decay of  $\varepsilon_k(\mathbf{W}_{q,\alpha}^r, L_p)$ . In the case  $d = 2$  the picture is almost complete. In the case  $d > 2$  the situation is fundamentally different. The problem of the right order of decay of  $\varepsilon_k(\mathbf{W}_{q,\alpha}^r, L_p)$  is still open for  $q = 1$ ,  $1 \leq p \leq \infty$  and  $p = \infty$ ,  $1 \leq q \leq \infty$ . In particular, it is open in the case  $q = 2$ ,  $p = \infty$ ,  $r = 1$  that is related to the small ball problem. We discuss in more detail the case  $p = \infty$ ,  $1 \leq q \leq \infty$ . We pointed out above that in the case  $d = 2$  the proof of lower bounds (the most difficult part) was based on the Small Ball Inequalities for the Haar system for  $r = 1$  and for the trigonometric system for all  $r$ . The existing conjecture is that

$$\varepsilon_k(\mathbf{W}_{q,\alpha}^r, L_\infty) \asymp k^{-r}(\ln k)^{(d-1)r+1/2}, \quad 1 < q < \infty, \quad (1.8)$$

for large enough  $r$ . The upper bound in (1.8) follows from (1.6). It is known that the corresponding lower bound in (1.8) would follow from the  $d$ -dimensional version of the Small Ball Inequality for the trigonometric system.

The main goal of this paper is to develop new techniques for proving upper bounds for the entropy numbers. We consider here slightly more general

classes than classes  $\mathbf{W}_{q,\alpha}^r$ . Let  $\mathbf{s} = (s_1, \dots, s_d)$  be a vector with nonnegative integer coordinates ( $\mathbf{s} \in \mathbb{Z}_+^d$ ) and

$$\rho(\mathbf{s}) := \{\mathbf{k} = (k_1, \dots, k_d) \in \mathbb{Z}_+^d : [2^{s_j-1}] \leq |k_j| < 2^{s_j}, \quad j = 1, \dots, d\}$$

where  $[a]$  denotes the integer part of a number  $a$ . Define for  $f \in L_1$

$$\delta_{\mathbf{s}}(f) := \sum_{\mathbf{k} \in \rho(\mathbf{s})} \hat{f}(\mathbf{k}) e^{i(\mathbf{k}, \mathbf{x})},$$

and

$$f_l := \sum_{\|\mathbf{s}\|_1=l} \delta_{\mathbf{s}}(f), \quad l \in \mathbb{N}_0, \quad \mathbb{N}_0 := \mathbb{N} \cup \{0\}.$$

Consider the class (see [33])

$$\mathbf{W}_q^{a,b} := \{f : \|f_l\|_q \leq 2^{-al} (\bar{l})^{(d-1)b}, \quad \bar{l} := \max(l, 1)\}.$$

Define

$$\|f\|_{\mathbf{W}_q^{a,b}} := \sup_l \|f_l\|_q 2^{al} (\bar{l})^{-(d-1)b}.$$

It is well known that the class  $\mathbf{W}_{q,\alpha}^r$  is embedded in the class  $\mathbf{W}_q^{r,0}$  for  $1 < q < \infty$ . Classes  $\mathbf{W}_q^{a,b}$  provide control of smoothness at two scales:  $a$  controls the power type smoothness and  $b$  controls the logarithmic scale smoothness. Similar classes with the power and logarithmic scales of smoothness are studied in the recent book of Triebel [34]. Here is one more class, which is equivalent to  $\mathbf{W}_q^{a,b}$  in the case  $1 < q < \infty$  (see [33]). Consider a class  $\bar{\mathbf{W}}_q^{a,b}$ , which consists of functions  $f$  with a representation (see Subsection 2.2 below for the definition of  $\mathcal{T}(Q_n)$ )

$$f = \sum_{n=1}^{\infty} t_n, \quad t_n \in \mathcal{T}(Q_n), \quad \|t_n\|_q \leq 2^{-an} n^{b(d-1)}.$$

In the case  $q = 1$  classes  $\bar{\mathbf{W}}_1^{a,b}$  are wider than  $\mathbf{W}_1^{a,b}$ .

The main results of the paper are the following theorems in the case  $d = 2$  for the extreme values of  $q = 1$  and  $q = \infty$ . First, we formulate two theorems for the case  $q = 1$ .

**Theorem 1.3.** *Let  $1 \leq p < \infty$  and  $a > 1 - 1/p$ . Then for  $d = 2$*

$$\varepsilon_k(\mathbf{W}_1^{a,b}, L_p) \asymp \varepsilon_k(\bar{\mathbf{W}}_1^{a,b}, L_p) \asymp k^{-a} (\log k)^{a+b}. \quad (1.9)$$

**Theorem 1.4.** *Let  $d = 2$  and  $a > 1$ . Then*

$$\varepsilon_k(\mathbf{W}_1^{a,b}, L_\infty) \asymp \varepsilon_k(\bar{\mathbf{W}}_1^{a,b}, L_\infty) \asymp k^{-a}(\log k)^{a+b+1/2}. \quad (1.10)$$

Second, we formulate three theorems for the case  $q = \infty$ .

**Theorem 1.5.** *We have for all  $d \geq 2$*

$$\varepsilon_k(\mathbf{W}_\infty^{a,b}, L_1) \gg k^{-a}(\log k)^{(d-1)(a+b)-1/2}. \quad (1.11)$$

**Theorem 1.6.** *We have for  $d = 2$ ,  $1 \leq p < \infty$ ,  $a \geq \max(1/2, 1 - 1/p)$*

$$\varepsilon_k(\mathbf{W}_\infty^{a,b}, L_p) \asymp \varepsilon_k(\bar{\mathbf{W}}_\infty^{a,b}, L_p) \asymp k^{-a}(\log k)^{a+b-1/2}. \quad (1.12)$$

**Theorem 1.7.** *We have for all  $d \geq 2$ ,  $1 \leq q \leq \infty$ ,  $a > 0$*

$$\varepsilon_k(\mathbf{W}_q^{a,b}, L_q) \asymp \varepsilon_k(\bar{\mathbf{W}}_q^{a,b}, L_q) \asymp k^{-a}(\log k)^{(d-1)(a+b)}. \quad (1.13)$$

Let us make some comments on Theorem 1.3. As we already mentioned above classes  $\mathbf{W}_1^{r,0}$  are close to classes  $\mathbf{W}_{1,\alpha}^r$  but they are different. We show that they are different even in the sense of asymptotic behavior of their entropy numbers. We point out that the right order of  $\varepsilon_k(\mathbf{W}_{1,\alpha}^r, L_p)$  is not known for  $d > 2$ . We confine ourselves to the case  $d = 2$ . It is proved in [11] that for  $r > \max(1/2, 1 - 1/p)$

$$\varepsilon_k(\mathbf{W}_{1,\alpha}^r, L_p) \asymp k^{-r}(\log k)^{r+1/2}, \quad 1 \leq p < \infty. \quad (1.14)$$

Theorem 1.3 gives for  $r > 1 - 1/p$

$$\varepsilon_k(\mathbf{W}_1^{r,0}, L_p) \asymp k^{-r}(\log k)^r, \quad 1 \leq p < \infty. \quad (1.15)$$

This shows that in the sense of the entropy numbers class  $\mathbf{W}_1^{r,0}$  is smaller than  $\mathbf{W}_{1,\alpha}^r$ . It is interesting to compare (1.14) and (1.15) with the known estimates in the case  $1 < q, p < \infty$

$$\varepsilon_k(\mathbf{W}_{q,\alpha}^r, L_p) \asymp \varepsilon_k(\mathbf{W}_q^{r,0}, L_p) \asymp k^{-r}(\log k)^r, \quad 1 \leq p < \infty. \quad (1.16)$$

Relation (1.16) is for the case  $d = 2$ . The general case of  $d$  is also known in this case (see (1.2) and its discussion above and also see Section 3.6 of [30] for the corresponding results and historical comments). Relations (1.15) and (1.16) show that in the sense of entropy numbers the class  $\mathbf{W}_1^{r,0}$  behaves as a limiting case of classes  $\mathbf{W}_{q,\alpha}^r$  when  $q \rightarrow 1$ .

The proof of upper bounds in Theorems 1.3 and 1.4 is based on greedy approximation technique. It is a new and powerful technique. In particular, Theorem 1.4 gives the same upper bound as in (1.6) for the class  $\mathbf{W}_1^{r,0}$ , which is wider than any of the classes  $\mathbf{W}_q^r$ ,  $q > 1$ , from (1.6). In Section 7 we develop mentioned above new technique, which is based on nonlinear  $m$ -term approximations, to prove the following result.

**Theorem 1.8.** *Let  $1 < q \leq 2$  and  $a > 1/q$ . Then*

$$\varepsilon_k(\mathbf{W}_q^{a,b}, L_\infty) \ll k^{-a}(\log k)^{(d-1)(a+b)+1/2}. \quad (1.17)$$

In particular, Theorem 1.8 implies (1.6).

Theorem 1.6 discovers an interesting new phenomenon. Comparing (1.12) with (1.13), we see that the entropy numbers of the class  $\mathbf{W}_\infty^{a,b}$  in the  $L_p$  space have different rate of decay in cases  $1 \leq p < \infty$  and  $p = \infty$ . We note that in the proof of the upper bounds in this new phenomenon we use the Riesz products for the hyperbolic crosses. This technique works well in the case  $d = 2$  but we do not know how to extend it to the general case  $d > 2$ . This difficulty is of the same nature as the corresponding difficulty in generalizing the Small Ball Inequality from  $d = 2$  to  $d > 2$  (see [30], Ch. 3, for further discussion). We already mentioned above that in studying the entropy numbers of function classes the discretization technique is useful. Classically, the Marcinkiewicz theorem serves as a powerful tool for discretizing the  $L_p$ -norm of a trigonometric polynomial. It works well in the multivariate case for trigonometric polynomials with frequencies from a parallelepiped. However, there is no analog of Marcinkiewicz' theorem for hyperbolic cross polynomials (see [11] and [5], Section 2.5, for a discussion). Thus, in Sections 5–7 we develop a new technique for estimating the entropy numbers of the unit balls of the hyperbolic cross polynomials. The most interesting results are obtained in the dimension  $d = 2$ . It would be very interesting to extend these results to the case  $d > 2$ . It is a challenging open problem.

Finally, we emphasize that in the case  $1 < q < \infty$ , when the classes  $\mathbf{W}_{q,\alpha}^r$  are embedded in the classes  $\mathbf{W}_q^{r,0}$ , the new technique, developed in this paper, provides all known upper bounds. In the case  $p < \infty$  Theorem 5.1 gives the upper bounds in (1.2), and in the case  $p = \infty$  Theorem 1.8 gives the upper bounds in (1.6).



## 2 Known results

### 2.1 General inequalities

For the reader's convenience we collect in this section known results, which will be used in this paper. The reader can find results of this subsection, except Theorem 2.3, and their proofs in [30], Chapter 3.

**Proposition 2.1.** *Let  $A \subset Y$ , and let  $Y$  be a subspace of  $X$ . Then*

$$N_\varepsilon(A, X) \geq N_{2\varepsilon}(A, Y).$$

Let us consider the space  $\mathbb{R}^D$  equipped with different norms, say, norms  $\|\cdot\|_X$  and  $\|\cdot\|_Y$ . For a Lebesgue measurable set  $E \in \mathbb{R}^D$  we denote its Lebesgue measure by  $\text{vol}(E) := \text{vol}_D(E)$ .

**Theorem 2.1.** *For any two norms  $X$  and  $Y$  and any  $\varepsilon > 0$  we have*

$$\frac{1}{\varepsilon^D} \frac{\text{vol}(B_Y)}{\text{vol}(B_X)} \leq N_\varepsilon(B_Y, X) \leq \frac{\text{vol}(B_Y(0, 2/\varepsilon) \oplus B_X)}{\text{vol}(B_X)}. \quad (2.1)$$

Let us formulate one immediate corollary of Theorem 2.1.

**Corollary 2.1.** *For any  $D$ -dimensional real Banach space  $X$  we have*

$$\varepsilon^{-D} \leq N_\varepsilon(B_X, X) \leq (1 + 2/\varepsilon)^D,$$

and, therefore,

$$\varepsilon_k(B_X, X) \leq 3(2^{-k/D}).$$

Let  $|\cdot| := \|\cdot\|_2$  denote the  $\ell_2^D$  norm and let  $B_2^D$  be a unit ball in  $\ell_2^D$ . Denote  $S^{D-1}$  the boundary of  $B_2^D$ . We define by  $d\sigma(x)$  the normalized  $(D-1)$ -dimensional measure on  $S^{D-1}$ . Consider another norm  $\|\cdot\|$  on  $\mathbb{R}^D$  and denote by  $X$  the  $\mathbb{R}^D$  equipped with  $\|\cdot\|$ .

**Theorem 2.2.** *Let  $X$  be  $\mathbb{R}^D$  equipped with  $\|\cdot\|$  and*

$$M_X := \int_{S^{D-1}} \|x\| d\sigma(x).$$

Then we have

$$\varepsilon_k(B_2^D, X) \ll M_X \begin{cases} (D/k)^{1/2}, & k \leq D \\ 2^{-k/D}, & k \geq D. \end{cases}$$

The following Nikol'skii-type inequalities are known (see [20], Chapter 1, Section 2).

**Theorem 2.3.** *Let  $1 \leq q < p < \infty$ . For any  $t \in \mathcal{T}(Q_n)$  (see Subsection 2.2 below for the definition of  $\mathcal{T}(Q_n)$ ) we have*

$$\|t\|_p \leq C(q, p, d) 2^{\beta n} \|t\|_q, \quad \beta := 1/q - 1/p.$$

## 2.2 Volume estimates

Denote for a natural number  $n$

$$Q_n := \cup_{\|\mathbf{s}\|_1 \leq n} \rho(\mathbf{s}); \quad \Delta Q_n := Q_n \setminus Q_{n-1} = \cup_{\|\mathbf{s}\|_1 = n} \rho(\mathbf{s})$$

with  $\|\mathbf{s}\|_1 = s_1 + \dots + s_d$  for  $\mathbf{s} \in \mathbb{Z}_+^d$ . We call a set  $\Delta Q_n$  *hyperbolic layer*. For a set  $\Lambda \subset \mathbb{Z}^d$  denote

$$\mathcal{T}(\Lambda) := \{f \in L_1 : \hat{f}(\mathbf{k}) = 0, \mathbf{k} \in \mathbb{Z}^d \setminus \Lambda\}, \quad \mathcal{T}(\Lambda)_p := \{f \in \mathcal{T}(\Lambda) : \|f\|_p \leq 1\}.$$

For a finite set  $\Lambda$  we assign to each  $f = \sum_{\mathbf{k} \in \Lambda} \hat{f}(\mathbf{k}) e^{i(\mathbf{k}, \mathbf{x})} \in \mathcal{T}(\Lambda)$  a vector

$$A(f) := \{(\operatorname{Re} \hat{f}(\mathbf{k}), \operatorname{Im} \hat{f}(\mathbf{k})), \quad \mathbf{k} \in \Lambda\} \in \mathbb{R}^{2|\Lambda|}$$

where  $|\Lambda|$  denotes the cardinality of  $\Lambda$  and define

$$B_\Lambda(L_p) := \{A(f) : f \in \mathcal{T}(\Lambda)_p\}.$$

The volume estimates of the sets  $B_\Lambda(L_p)$  and related questions have been studied in a number of papers: the case  $\Lambda = [-n, n]$ ,  $p = \infty$  in [8]; the case  $\Lambda = [-N_1, N_1] \times \dots \times [-N_d, N_d]$ ,  $p = \infty$  in [22], [23]. In the case  $\Lambda = \Pi(\mathbf{N}, d) := [-N_1, N_1] \times \dots \times [-N_d, N_d]$ ,  $\mathbf{N} := (N_1, \dots, N_d)$ , the following estimates are known.

**Theorem 2.4.** *For any  $1 \leq p \leq \infty$  we have*

$$(\operatorname{vol}(B_{\Pi(\mathbf{N}, d)}(L_p)))^{(2|\Pi(\mathbf{N}, d)|)^{-1}} \asymp |\Pi(\mathbf{N}, d)|^{-1/2},$$

with constants in  $\asymp$  that may depend only on  $d$ .

We note that the most difficult part of Theorem 2.4 is the lower estimate for  $p = \infty$ . The corresponding estimate was proved in the case  $d = 1$  in [8] and in the general case in [22] and [23] by a method different from the one in [8]. The upper estimate for  $p = 1$  in Theorem 2.4 can be easily reduced to the volume estimate for an octahedron (see, for instance [24]). In the case  $p = 2$  Theorem 2.4 is a direct corollary of the well known estimates of the volume of the Euclidean unit ball.

The case of arbitrary  $\Lambda$  and  $p = 1$  was studied in [9]. The results of [9] imply the following estimate.

**Theorem 2.5.** *For any finite set  $\Lambda \subset \mathbb{Z}^d$  and any  $1 \leq p \leq 2$  we have*

$$\text{vol}(B_\Lambda(L_p))^{(2|\Lambda|)^{-1}} \asymp |\Lambda|^{-1/2}.$$

The following result was obtained in [11].

**Theorem 2.6.** *Let  $\Lambda$  have the form  $\Lambda = \cup_{\mathbf{s} \in S} \rho(\mathbf{s})$ ,  $S \subset \mathbb{Z}_+^d$  is a finite set. Then for any  $1 \leq p < \infty$  we have*

$$\text{vol}(B_\Lambda(L_p))^{(2|\Lambda|)^{-1}} \asymp |\Lambda|^{-1/2}.$$

In particular, Theorem 2.6 implies for  $d = 2$  and  $1 \leq p < \infty$  that

$$(\text{vol}(B_{\Delta Q_n}(L_p)))^{(2|\Delta Q_n|)^{-1}} \asymp |\Delta Q_n|^{-1/2} \asymp (2^n n)^{-1/2}. \quad (2.2)$$

The following result was obtained in [11]. Denote  $D := 2|\Delta Q_n|$ .

**Theorem 2.7.** *In the case  $d = 2$  we have*

$$(\text{vol}(B_{\Delta Q_n}(L_\infty)))^{1/D} \asymp (2^n n^2)^{-1/2}. \quad (2.3)$$

The following lemma from [11] is an important ingredient of analysis in this paper. For the reader's convenience we give a proof of this lemma here.

**Lemma 2.1.** *Let  $\Lambda \subseteq [-2^n, 2^n]^d$  and  $N := 2|\Lambda|$ . Then*

$$(\text{vol}(B_\Lambda(L_\infty)))^{1/N} \geq C(d)(Nn)^{-1/2}.$$

*Proof.* We use the following result of E. Gluskin [6].

**Theorem 2.8.** *Let  $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_M\} \subset \mathbb{R}^N$ ,  $\|\mathbf{y}_i\| = 1$ ,  $i = 1, \dots, M$  and*

$$W(Y) := \{\mathbf{x} \in \mathbb{R}^N : |(\mathbf{x}, \mathbf{y}_i)| \leq 1, \quad i = 1, \dots, M\}.$$

*Then*

$$(\text{vol}(W(Y)))^{1/N} \geq C(1 + \log(M/N))^{-1/2}.$$

Consider the following lattice on the  $\mathbb{T}^d$ :

$$G_n := \{\mathbf{x}(l) = (l_1, \dots, l_d)\pi 2^{-n-1}, \quad 1 \leq l_j \leq 2^{n+2}, \quad l_j \in \mathbb{N}, \quad j = 1, \dots, d\}.$$

It is clear that  $|G_n| = 2^{d(n+2)}$ . It is well known (see [25], Ch.2, Theorem 2.4) that for any  $f \in \mathcal{T}([-2^n, 2^n]^d)$  one has

$$\|f\|_\infty \leq C_1(d) \max_{\mathbf{x} \in G_n} |f(\mathbf{x})|.$$

Thus, for any  $\Lambda \subseteq [-2^n, 2^n]^d$  we have

$$\{A(f) : f \in \mathcal{T}(\Lambda), \quad |f(\mathbf{x})| \leq C_1(d)^{-1}, \quad \mathbf{x} \in G_n\} \subseteq B_\Lambda(L_\infty). \quad (2.4)$$

Further

$$\begin{aligned} |f(\mathbf{x})|^2 &= \left| \sum_{\mathbf{k} \in \Lambda} \hat{f}(\mathbf{k}) e^{i(\mathbf{k}, \mathbf{x})} \right|^2 = \\ &= \left( \sum_{\mathbf{k} \in \Lambda} \text{Re} \hat{f}(\mathbf{k}) \cos(\mathbf{k}, \mathbf{x}) - \text{Im} \hat{f}(\mathbf{k}) \sin(\mathbf{k}, \mathbf{x}) \right)^2 \\ &+ \left( \sum_{\mathbf{k} \in \Lambda} \text{Re} \hat{f}(\mathbf{k}) \sin(\mathbf{k}, \mathbf{x}) + \text{Im} \hat{f}(\mathbf{k}) \cos(\mathbf{k}, \mathbf{x}) \right)^2. \end{aligned}$$

We associate with each point  $\mathbf{x} \in G_n$  two vectors  $\mathbf{y}^1(\mathbf{x})$  and  $\mathbf{y}^2(\mathbf{x})$  from  $\mathbb{R}^N$ :

$$\mathbf{y}^1(\mathbf{x}) := \{(\cos(\mathbf{k}, \mathbf{x}), -\sin(\mathbf{k}, \mathbf{x})), \quad \mathbf{k} \in \Lambda\},$$

$$\mathbf{y}^2(\mathbf{x}) := \{(\sin(\mathbf{k}, \mathbf{x}), \cos(\mathbf{k}, \mathbf{x})), \quad \mathbf{k} \in \Lambda\}.$$

Then

$$\|\mathbf{y}^1(\mathbf{x})\|^2 = \|\mathbf{y}^2(\mathbf{x})\|^2 = |\Lambda|$$

and

$$|f(\mathbf{x})|^2 = (A(f), \mathbf{y}^1(\mathbf{x}))^2 + (A(f), \mathbf{y}^2(\mathbf{x}))^2.$$

It is clear that the condition  $|f(\mathbf{x})| \leq C_1(d)^{-1}$  is satisfied if

$$|(A(f), \mathbf{y}^i(\mathbf{x}))| \leq 2^{-1/2} C_1(d)^{-1}, \quad i = 1, 2.$$

Let now

$$Y := \{\mathbf{y}^i(\mathbf{x}) / \|\mathbf{y}^i(\mathbf{x})\|, \quad \mathbf{x} \in G_n, \quad i = 1, 2\}.$$

Then  $M = 2^{d(n+2)+1}$  and by Theorem 2.8

$$(\text{vol}(W(Y)))^{1/N} \gg (1 + \log(M/N))^{-1/2} \gg n^{-1/2}. \quad (2.5)$$

Using that the condition

$$|(A(f), \mathbf{y}^i(\mathbf{x}))| \leq 1$$

is equivalent to the condition

$$|(A(f), \mathbf{y}^i(\mathbf{x}) / \|\mathbf{y}^i(\mathbf{x})\|)| \leq (N/2)^{-1/2}$$

we get from (2.4) and (2.5)

$$(\text{vol}(B_\Lambda(L_\infty)))^{1/N} \gg (Nn)^{-1/2}.$$

This completes the proof of Lemma 2.1 □

### 3 New lower bounds. The volumes technique

In this section we prove lower bounds in Theorems 1.3 – 1.7 from the Introduction.

**Proof of lower bounds in Theorems 1.3 and 1.7.** The lower bound in Theorem 1.3 follows from the lower bound in Theorem 1.7 with  $q = 1$ . We prove the lower bounds for the  $\varepsilon_k(\mathbf{W}_q^{a,b}, L_q)$  with  $1 \leq q \leq \infty$  and any  $d$ . This lower bound is derived from the well known simple inequality (see Corollary 2.1 above)

$$N_\varepsilon(B_X, X) \geq \varepsilon^{-D} \quad (3.1)$$

for any  $D$ -dimensional real Banach space  $X$ . Consider as a Banach space  $X$  the  $\mathcal{T}(\Delta Q_n)$  with  $L_q$  norm. Clearly, it can be seen as a  $D$ -dimensional real Banach space with  $D = 2|\Delta Q_n|$ . It follows from the definition of  $\mathbf{W}_q^{a,b}$  that

$$2^{-an} n^{b(d-1)} \mathcal{T}(\Delta Q_n)_q \subset \mathbf{W}_q^{a,b}. \quad (3.2)$$

Take  $k = 2|\Delta Q_n|$ . Then (3.1) implies that

$$\varepsilon_k(\mathcal{T}(\Delta Q_n)_q, L_q \cap \mathcal{T}(\Delta Q_n)) \gg 1. \quad (3.3)$$

We now use one more well known fact from the entropy theory – Proposition 2.1. This and inequality (3.3) imply

$$\varepsilon_k(\mathcal{T}(\Delta Q_n)_q, L_q) \gg 1. \quad (3.4)$$

Taking into account (3.2) and the fact  $k \asymp 2^n n^{d-1}$  we derive from (3.4) the required lower bound for the  $\varepsilon_k(\mathbf{W}_q^{a,b}, L_q)$ .

The lower bounds in Theorems 1.3 and 1.7 are proved.

**Proof of lower bounds in Theorem 1.4.** We prove the lower bound for  $\varepsilon_k(\mathbf{W}_1^{a,b}, L_\infty)$ . This proof is somewhat similar to the proof of lower bounds in Theorem 1.3. Instead of (3.1) we now use the inequality (see Theorem 2.1 above)

$$N_\varepsilon(B_Y, X) \geq \varepsilon^{-D} \frac{\text{vol}(B_Y)}{\text{vol}(B_X)} \quad (3.5)$$

with  $B_Y := B_{\Delta Q_n}(L_1)$  and  $B_X := B_{\Delta Q_n}(L_\infty)$ . It follows from the definition of  $\mathbf{W}_1^{a,b}$  that

$$2^{-an} n^{b(d-1)} \mathcal{T}(\Delta Q_n)_1 \subset \mathbf{W}_1^{a,b}. \quad (3.6)$$

Take  $k = 2|\Delta Q_n|$ . Then (3.5), Theorem 2.7, and (2.2) imply that

$$\varepsilon_k(\mathcal{T}(\Delta Q_n)_1, L_\infty \cap \mathcal{T}(\Delta Q_n)) \gg n^{1/2}. \quad (3.7)$$

Proposition 2.1 and inequality (3.7) imply

$$\varepsilon_k(\mathcal{T}(\Delta Q_n)_1, L_\infty) \gg n^{1/2}. \quad (3.8)$$

Taking into account (3.6) and the fact  $k \asymp 2^n n^{d-1}$  we derive from (3.8) the required lower bound for the  $\varepsilon_k(\mathbf{W}_1^{a,b}, L_\infty)$ .

The lower bounds in Theorem 1.4 are proved.

**Proof of Theorem 1.5.** We prove the lower bound for  $\varepsilon_k(\mathbf{W}_\infty^{a,b}, L_1)$ . This proof goes along the lines of the above proof of lower bounds in Theorem 1.4. We use (3.5) with  $B_X := B_{\Delta Q_n}(L_1)$  and  $B_Y := B_{\Delta Q_n}(L_\infty)$ . It follows from the definition of  $\mathbf{W}_\infty^{a,b}$  that

$$2^{-an} n^{b(d-1)} \mathcal{T}(\Delta Q_n)_\infty \subset \mathbf{W}_\infty^{a,b}. \quad (3.9)$$

Take  $k = 2|\Delta Q_n|$ . Then (3.5), Lemma 2.1 with  $\Lambda = \Delta Q_n$ , and Theorem 2.5 with  $\Lambda = \Delta Q_n$  imply that

$$\varepsilon_k(\mathcal{T}(\Delta Q_n)_\infty, L_1 \cap \mathcal{T}(\Delta Q_n)) \gg n^{-1/2}. \quad (3.10)$$

Proposition 2.1 and inequality (3.10) imply

$$\varepsilon_k(\mathcal{T}(\Delta Q_n)_\infty, L_1) \gg n^{-1/2}. \quad (3.11)$$

Taking into account (3.9) and the fact  $k \asymp 2^n n^{d-1}$  we derive from (3.11) the required lower bound for the  $\varepsilon_k(\mathbf{W}_\infty^{a,b}, L_1)$ .

The lower bounds in Theorem 1.5 are proved.

**Proof of lower bounds in Theorem 1.6.** The required lower bounds follow from Theorem 1.5.

## 4 Upper bounds. A general scheme

**From finite dimensional to infinite dimensional.** Let  $X$  and  $Y$  be two Banach spaces. We discuss a problem of estimating the entropy numbers of an approximation class, defined in the space  $X$ , in the norm of the space  $Y$ . Suppose a sequence of finite dimensional subspaces  $X_n \subset X$ ,  $n = 1, \dots$ , is given. Define the following class

$$\bar{\mathbf{W}}_X^{a,b} := \bar{\mathbf{W}}_X^{a,b}\{X_n\} := \{f \in X : f = \sum_{n=1}^{\infty} f_n, \quad f_n \in X_n,$$

$$\|f_n\|_X \leq 2^{-an} n^b, \quad n = 1, 2, \dots\}.$$

In particular,

$$\bar{\mathbf{W}}_q^{a,b} = \bar{\mathbf{W}}_{L_q}^{a,b(d-1)}\{\mathcal{T}(Q_n)\}.$$

Denote  $D_n := \dim X_n$  and assume that for the unit balls  $B(X_n) := \{f \in X_n : \|f\|_X \leq 1\}$  we have the following upper bounds for the entropy numbers: there exist real  $\alpha$  and nonnegative  $\gamma$  and  $\beta \in (0, 1]$  such that

$$\varepsilon_k(B(X_n), Y) \ll n^\alpha \begin{cases} (D_n/(k+1))^\beta (\log(4D_n/(k+1)))^\gamma, & k \leq 2D_n, \\ 2^{-k/(2D_n)}, & k \geq 2D_n. \end{cases} \quad (4.1)$$

**Theorem 4.1.** Assume  $D_n \asymp 2^n n^c$ ,  $c \geq 0$ ,  $a > \beta$ , and subspaces  $\{X_n\}$  satisfy (4.1). Then

$$\varepsilon_k(\bar{\mathbf{W}}_X^{a,b}\{X_n\}, Y) \ll k^{-a}(\log k)^{ac+b+\alpha}. \quad (4.2)$$

*Proof.* For a given  $k$  let  $n$  be such that  $k \asymp D_n \asymp 2^n n^c$ . It follows from the definition of class  $\bar{\mathbf{W}}_X^{a,b}$  that

$$\varepsilon_k(\bar{\mathbf{W}}_X^{a,b}, Y) \leq \sum_{l=1}^{\infty} 2^{-al} l^b \varepsilon_{k_l}(B(X_l), Y),$$

provided  $\sum_{l=1}^{\infty} k_l \leq k$ . For  $l < n$  we define  $k_l := \lceil 3a(n-l)D_l/\beta \rceil$ . Then  $\sum_{l=1}^{n-1} k_l \ll k$  and by our assumption (4.1)

$$\begin{aligned} & \sum_{l=1}^{n-1} 2^{-al} l^b \varepsilon_{k_l}(B(X_l), Y) \\ & \ll \sum_{l=1}^n 2^{-al} l^{b+\alpha} 2^{-k_l/(2D_l)} \\ & \ll 2^{-an} n^{b+\alpha} \ll k^{-a}(\log k)^{ac+b+\alpha}. \end{aligned}$$

For  $l \geq n$  we define  $k_l := \lceil D_n 2^{\mu(n-l)} \rceil$ ,  $\mu := (a - \beta)/(2\beta)$ . Then  $\sum_{l \geq n} k_l \ll k$ . Therefore, by our assumption (4.1) we get

$$\begin{aligned} & \sum_{l \geq n} 2^{-al} l^b \varepsilon_{k_l}(B(X_l), Y) \\ & \ll \sum_{l \geq n} 2^{-al} l^{b+\alpha} 2^{\mu(l-n)\beta} (D_l/D_n)^{\beta} (l-n)^{\gamma} \\ & \ll 2^{-an} n^{b+\alpha} \ll k^{-a}(\log k)^{ac+b+\alpha}. \end{aligned}$$

Thus we proved

$$\varepsilon_{Ck}(\bar{\mathbf{W}}_X^{a,b}, Y) \ll k^{-a}(\log k)^{ac+b+\alpha}. \quad (4.3)$$

Taking into account that the right hand side in (4.3) decays polynomially, we conclude that the upper bound in (4.2) holds.  $\square$

**Remark 4.1.** In the case  $Y = X$  Theorem 4.1 holds without assumption (4.1). It is sufficient to use Corollary 2.1.



**From  $m$ -term approximation to entropy numbers.** We discuss a technique, which is based on the following two steps strategy. At the first step we obtain bounds of the best  $m$ -term approximations with respect to a dictionary. At the second step we use general inequalities relating the entropy numbers to the best  $m$ -term approximations. We begin the detailed discussion with the second step of the above strategy. Let  $\mathcal{D} = \{g_j\}_{j=1}^N$  be a system of elements of cardinality  $|\mathcal{D}| = N$  in a Banach space  $X$ . Consider best  $m$ -term approximations of  $f$  with respect to  $\mathcal{D}$

$$\sigma_m(f, \mathcal{D})_X := \inf_{\{c_j\}; \Lambda: |\Lambda|=m} \|f - \sum_{j \in \Lambda} c_j g_j\|.$$

For a function class  $F$  set

$$\sigma_m(F, \mathcal{D})_X := \sup_{f \in F} \sigma_m(f, \mathcal{D})_X.$$

The following results are from [31].

**Theorem 4.2.** *Let a compact  $F \subset X$  be such that there exists a system  $\mathcal{D}$ ,  $|\mathcal{D}| = N$ , and a number  $r > 0$  such that*

$$\sigma_m(F, \mathcal{D})_X \leq m^{-r}, \quad m \leq N.$$

*Then for  $k \leq N$*

$$\varepsilon_k(F, X) \leq C(r) \left( \frac{\log(2N/k)}{k} \right)^r. \quad (4.4)$$

**Remark 4.2.** *Suppose that a compact  $F$  from Theorem 4.2 belongs to an  $N$ -dimensional subspace  $X_N := \text{span}(\mathcal{D})$ . Then in addition to (4.4) we have for  $k \geq N$*

$$\varepsilon_k(F, X) \leq C(r) N^{-r} 2^{-k/(2N)}. \quad (4.5)$$

We point out that Remark 4.2 is formulated for a complex Banach space  $X$ . In the case of real Banach space  $X$  we have  $2^{-k/N}$  instead of  $2^{-k/(2N)}$  in (4.5).

## 5 Hyperbolic cross polynomials, $1 < q < p < \infty$

$\infty$

We now proceed to the first step of the above described strategy. First, we discuss the entropy numbers  $\varepsilon_k(\mathcal{T}(Q_n)_q, L_p)$  in the case  $1 < q \leq 2 \leq p < \infty$ .

For the  $m$ -term approximation we use the following system described and studied in [29]. We define a system of orthogonal trigonometric polynomials which is optimal in a certain sense (see [29]) for  $m$ -term approximations. Variants of this system are well-known and very useful in interpolation of functions by trigonometric polynomials. We begin with a construction of the system  $\mathcal{U}$  in the univariate case. Denote

$$U_n^+(x) := \sum_{k=0}^{2^n-1} e^{ikx} = \frac{e^{i2^n x} - 1}{e^{ix} - 1}, \quad n = 0, 1, 2, \dots;$$

$$U_{n,j}^+(x) := e^{i2^n x} U_n^+(x - 2\pi j 2^{-n}), \quad j = 0, 1, \dots, 2^n - 1;$$

$$U_{n,j}^-(x) := e^{-i2^n x} U_n^+(-x + 2\pi j 2^{-n}), \quad j = 0, 1, \dots, 2^n - 1.$$

It will be more convenient for us to normalize in  $L_2$  the system of functions  $\{U_{n,j}^+, U_{n,j}^-\}$ . We write

$$u_{n,j}^+(x) := 2^{-n/2} U_{n,j}^+(x), \quad u_{n,j}^-(x) := 2^{-n/2} U_{n,j}^-(x).$$

For  $k = 2^n + j$ ,  $n = 0, 1, 2, \dots$ , and  $j = 0, 1, \dots, 2^n - 1$ , define

$$u_k(x) := u_{n,j}^+(x), \quad u_{-k}(x) := u_{n,j}^-(x).$$

The above formulas define  $u_k$  for all  $k \in \mathbb{Z} \setminus \{0\}$ . Finally, define  $u_0(x) = 1$ . Set  $\mathcal{U} := \{u_k\}_{k \in \mathbb{Z}}$ . In the multivariate case of  $\mathbf{x} = (x_1, \dots, x_d)$  we define the system  $\mathcal{U}^d$  as the tensor product of the univariate systems  $\mathcal{U}$ . Namely,  $\mathcal{U}^d := \{u_{\mathbf{k}}(\mathbf{x})\}_{\mathbf{k} \in \mathbb{Z}^d}$ , where

$$u_{\mathbf{k}}(\mathbf{x}) := \prod_{i=1}^d u_{k_i}(x_i), \quad \mathbf{k} = (k_1, \dots, k_d).$$

For  $s \in \mathbb{N}$  denote

$$\rho^+(s) := \{k : 2^{s-1} \leq k < 2^s\}, \quad \rho^-(s) := \{-k : 2^{s-1} \leq k < 2^s\},$$

and for  $s = 0$  denote

$$\rho^+(0) = \rho^-(0) = \rho(0) := \{0\}.$$

Then for  $\epsilon = +$  or  $\epsilon = -$  we have

$$\mathcal{T}(\rho^\epsilon(s)) = \text{span}\{u_k, \quad k \in \rho^\epsilon(s)\} = \text{span}\{e^{ikx}, \quad k \in \rho^\epsilon(s)\}.$$

In the multivariate case for  $\mathbf{s} = (s_1, \dots, s_d)$  and  $\epsilon = (\epsilon_1, \dots, \epsilon_d)$  denote

$$\rho^\epsilon(\mathbf{s}) := \rho^{\epsilon_1}(s_1) \times \dots \times \rho^{\epsilon_d}(s_d).$$

Then

$$\mathcal{T}(\rho^\epsilon(\mathbf{s})) = \text{span}\{u_{\mathbf{k}}, \quad \mathbf{k} \in \rho^\epsilon(\mathbf{s})\} = \text{span}\{e^{i(\mathbf{k}, \mathbf{x})}, \quad \mathbf{k} \in \rho^\epsilon(\mathbf{s})\}.$$

It is easy to check that for any  $\mathbf{k} \neq \mathbf{m}$  we have

$$\langle u_{\mathbf{k}}, u_{\mathbf{m}} \rangle = (2\pi)^{-d} \int_{\mathbb{T}^d} u_{\mathbf{k}}(\mathbf{x}) \bar{u}_{\mathbf{m}}(\mathbf{x}) d\mathbf{x} = 0,$$

and

$$\|u_{\mathbf{k}}\|_2 = 1.$$

We use the notations for  $f \in L_1$

$$f_{\mathbf{k}} := \langle f, u_{\mathbf{k}} \rangle := (2\pi)^{-d} \int_{\mathbb{T}^d} f(\mathbf{x}) \bar{u}_{\mathbf{k}}(\mathbf{x}) d\mathbf{x}; \quad \hat{f}(\mathbf{k}) := (2\pi)^{-d} \int_{\mathbb{T}^d} f(\mathbf{x}) e^{-i(\mathbf{k}, \mathbf{x})} d\mathbf{x}$$

and

$$\delta_{\mathbf{s}}^\epsilon(f) := \sum_{\mathbf{k} \in \rho^\epsilon(\mathbf{s})} \hat{f}(\mathbf{k}) e^{i(\mathbf{k}, \mathbf{x})}.$$

The following important for us analog of the Marcinkiewicz theorem holds

$$\|\delta_{\mathbf{s}}^\epsilon(f)\|_p^p \asymp \sum_{\mathbf{k} \in \rho^\epsilon(\mathbf{s})} \|f_{\mathbf{k}} u_{\mathbf{k}}\|_p^p, \quad 1 < p < \infty, \quad (5.1)$$

with constants depending on  $p$  and  $d$ .

We will often use the following inequalities

$$\left( \sum_{\mathbf{s}, \epsilon} \|\delta_{\mathbf{s}}^\epsilon(f)\|_p^p \right)^{1/p} \ll \|f\|_p \ll \left( \sum_{\mathbf{s}, \epsilon} \|\delta_{\mathbf{s}}^\epsilon(f)\|_p^2 \right)^{1/2}, \quad 2 \leq p < \infty, \quad (5.2)$$

$$\left( \sum_{\mathbf{s}, \epsilon} \|\delta_{\mathbf{s}}^\epsilon(f)\|_p^2 \right)^{1/2} \ll \|f\|_p \ll \left( \sum_{\mathbf{s}, \epsilon} \|\delta_{\mathbf{s}}^\epsilon(f)\|_p^p \right)^{1/p}, \quad 1 < p \leq 2, \quad (5.3)$$

which are corollaries of the well-known Littlewood-Paley inequalities

$$\|f\|_p \asymp \left\| \left( \sum_{\mathbf{s}} \left| \sum_{\epsilon} \delta_{\mathbf{s}}^\epsilon(f) \right|^2 \right)^{1/2} \right\|_p, \quad 1 < p < \infty. \quad (5.4)$$

**Lemma 5.1.** *Let  $1 < q \leq 2 \leq p < \infty$ . Let  $\mathcal{D}_n^1 := \{u_{\mathbf{k}} : \mathbf{k} \in \Delta Q_n\}$ . Then*

$$\sigma_m(\mathcal{T}(\Delta Q_n)_q, \mathcal{D}_n^1)_p \ll (|\Delta Q_n|/m)^\beta, \quad \beta = 1/q - 1/p. \quad (5.5)$$

*Proof.* Theorem 2.3 implies Lemma 5.1 for  $m \leq m_n := \lfloor |Q_n|2^{-n} \rfloor$ . Let  $m \geq m_n$ . Take  $f \in \mathcal{T}(\Delta Q_n)$ . Then

$$f = \sum_{\|\mathbf{s}\|_1=n} \sum_{\epsilon} \delta_{\mathbf{s}}^\epsilon(f).$$

Represent

$$\delta_{\mathbf{s}}^\epsilon(f) = \sum_{\mathbf{k} \in \rho^\epsilon(\mathbf{s})} f_{\mathbf{k}} u_{\mathbf{k}}.$$

For convenience we will omit  $\epsilon$  in the notations  $\delta_{\mathbf{s}}^\epsilon(f)$ ,  $\rho^\epsilon(\mathbf{s})$ , meaning that we are estimating a quantity  $\delta_{\mathbf{s}}^\epsilon(f)$  for a fixed  $\epsilon$  and all estimates we are going to do in this paper are the same for all  $\epsilon$ . We now need the following well known simple lemma (see, for instance, [20], p.92).

**Lemma 5.2.** *Let  $a_1 \geq a_2 \geq \dots \geq a_M \geq 0$  and  $1 \leq q \leq p \leq \infty$ . Then for all  $m < M$  one has*

$$\left( \sum_{k=m}^M a_k^p \right)^{1/p} \leq m^{-\beta} \left( \sum_{k=1}^M a_k^q \right)^{1/q}, \quad \beta := 1/q - 1/p.$$

We apply this lemma to each set of  $f_{\mathbf{k}}$ ,  $\mathbf{k} \in \rho(\mathbf{s})$ ,  $\|\mathbf{s}\|_1 = n$  with  $m_{\mathbf{s}} := \lfloor m/m_n \rfloor$ . Denote  $G_{\mathbf{s}}$  the set of cardinality  $|G_{\mathbf{s}}| = m_{\mathbf{s}}$  of  $\mathbf{k}$  from  $\rho(\mathbf{s})$  with largest  $|f_{\mathbf{k}}|$ . Then by (5.1) we obtain

$$\left\| \sum_{\mathbf{k} \in \rho(\mathbf{s}) \setminus G_{\mathbf{s}}} f_{\mathbf{k}} u_{\mathbf{k}} \right\|_p \asymp 2^{n(1/2-1/p)} \left( \sum_{\mathbf{k} \in \rho(\mathbf{s}) \setminus G_{\mathbf{s}}} |f_{\mathbf{k}}|^p \right)^{1/p}.$$

Applying Lemma 5.2 we continue

$$\ll 2^{n(1/2-1/p)} (m_{\mathbf{s}} + 1)^{-\beta} \left( \sum_{\mathbf{k} \in \rho(\mathbf{s})} |f_{\mathbf{k}}|^q \right)^{1/q}.$$

Using (5.1) again we obtain

$$\left\| \sum_{\mathbf{k} \in \rho(\mathbf{s}) \setminus G_{\mathbf{s}}} f_{\mathbf{k}} u_{\mathbf{k}} \right\|_p \ll (|\Delta Q_n|/m)^\beta \|\delta_{\mathbf{s}}(f)\|_q. \quad (5.6)$$

Estimating the  $\|\cdot\|_p$  from above by (5.2) and the  $\|\cdot\|_q$  from below by (5.3) we complete the proof of Lemma 5.1.  $\square$

It is easy to see that Lemma 5.1 implies the corresponding result for  $\mathcal{T}(Q_n)$ .

**Lemma 5.3.** *Let  $1 < q \leq 2 \leq p < \infty$ . Let  $\mathcal{D}'_n := \{u_{\mathbf{k}} : \mathbf{k} \in Q_n\}$ . Then*

$$\sigma_m(\mathcal{T}(Q_n)_q, \mathcal{D}'_n)_p \ll (|Q_n|/m)^\beta, \quad \beta = 1/q - 1/p. \quad (5.7)$$

We now apply the second step of the strategy described in Section 4. Theorem 4.2, Remark 4.2 and Lemma 5.3 imply the following lemma.

**Lemma 5.4.** *Let  $1 < q \leq 2 \leq p < \infty$  and  $\beta := 1/q - 1/p$ . Then*

$$\varepsilon_k(\mathcal{T}(Q_n)_q, L_p) \ll \begin{cases} (|Q_n|/k)^\beta (\log(4|Q_n|/k))^\beta, & k \leq 2|Q_n|, \\ 2^{-k/(2|Q_n|)}, & k \geq 2|Q_n|. \end{cases} \quad (5.8)$$

We now extend Lemma 5.4 to the case  $1 < q < p < \infty$ . We will use a decomposition technique.

**Lemma 5.5.** *Let  $1 \leq u < q < p \leq \infty$ . For any  $f \in L_q(\Omega)$ ,  $\|f\|_q \leq 1$ , and any positive numbers  $a, b$  there exists a representation  $f = f_1 + f_2$  such that*

$$a\|f_1\|_u \leq a^{1-\theta}b^\theta, \quad b\|f_2\|_p \leq a^{1-\theta}b^\theta, \quad \theta := (1/u - 1/q)(1/u - 1/p)^{-1}. \quad (5.9)$$

*Proof.* Let  $f_T$  denote the  $T$  cut off of  $f$ :  $f_T(x) = f(x)$  if  $|f(x)| \leq T$  and  $f_T(x) = 0$  otherwise. Clearly,  $\|f_T\|_\infty \leq T$ . Set  $f^T := f - f_T$ . We now estimate the  $L_u$  norm of the  $f^T$ . Let  $E := \{x : f^T(x) \neq 0\}$ . First, we bound from above the measure  $|E|$  of  $E$ . We have

$$1 \geq \int_\Omega |f(\mathbf{x})|^q d\mathbf{x} \geq \int_E T^q d\mathbf{x} = T^q |E|.$$

Second, we bound the  $\|f^T\|_u$

$$\|f^T\|_u^u = \int_E |f^T(\mathbf{x})|^u d\mathbf{x} \leq \left( \int_E |f^T(\mathbf{x})|^q d\mathbf{x} \right)^{u/q} |E|^{1-u/q} \leq T^{u-q}.$$

Third, we bound the  $\|f_T\|_p$ . Using the inequality

$$\|g\|_p \leq \|g\|_q^{q/p} \|g\|_\infty^{1-q/p}$$

we obtain

$$\|f_T\|_p \leq T^{1-q/p}.$$

Specifying  $T = (a/b)^{(q(1/u-1/p))^{-1}}$  we get

$$a\|f^T\|_u \leq a^{1-\theta}b^\theta, \quad b\|f_T\|_p \leq a^{1-\theta}b^\theta.$$

This proves the lemma.  $\square$

We now derive from Lemma 5.5 the following inequality for the entropy numbers.

**Lemma 5.6.** *For  $1 < u < q < p < \infty$  we have for  $\theta := (1/u - 1/q)(1/u - 1/p)^{-1}$*

$$\varepsilon_{k+l}(\mathcal{T}(Q_n)_q, L_p) \leq C(u, p, d) (\varepsilon_k(\mathcal{T}(Q_n)_u, L_p))^{1-\theta} (\varepsilon_l(\mathcal{T}(Q_n)_p, L_p))^\theta.$$

*Proof.* Let  $t \in \mathcal{T}(Q_n)_q$ . Applying Lemma 5.5 we split the polynomial  $t$  into a sum  $t = f_1 + f_2$  satisfying (5.9). Consider

$$t_1 := S_{Q_n}(f_1), \quad t_2 := S_{Q_n}(f_2).$$

Then

$$at_1 \in \mathcal{T}(Q_n)_u C(u, d) a^{1-\theta} b^\theta \quad \text{and} \quad bt_2 \in \mathcal{T}(Q_n)_p C(p, d) a^{1-\theta} b^\theta.$$

Let  $a$  and  $b$  be such that

$$a > \varepsilon_k(\mathcal{T}(Q_n)_u, L_p), \quad b > \varepsilon_l(\mathcal{T}(Q_n)_p, L_p).$$

Find  $y_1, \dots, y_{2^k}$  and  $z_1, \dots, z_{2^l}$  such that

$$\mathcal{T}(Q_n)_u \subset \cup_{i=1}^{2^k} B_{L_p}(y_i, a), \quad \mathcal{T}(Q_n)_p \subset \cup_{j=1}^{2^l} B_{L_p}(z_j, b).$$

Take any  $t \in \mathcal{T}(Q_n)_q$ . Set  $\varepsilon := a^{1-\theta}b^\theta \max\{C(u, d), C(p, d)\}$  and as above find  $t_1$  and  $t_2$  from  $\mathcal{T}(Q_n)$  such that  $t = t_1 + t_2$  and

$$a\|t_1\|_u \leq \varepsilon, \quad b\|t_2\|_p \leq \varepsilon.$$

Clearly, for some  $i$

$$at_1/\varepsilon \in B_{L_p}(y_i, a) \quad \Rightarrow \quad t_1 \in B_{L_p}(\varepsilon y_i/a, \varepsilon) \quad (5.10)$$

and for some  $j$

$$bt_2/\varepsilon \in B_{L_p}(z_j, b) \Rightarrow t_2 \in B_{L_p}(\varepsilon z_j/b, \varepsilon). \quad (5.11)$$

Consider the sets  $G_{i,j} := B_{L_p}(\varepsilon y_i/a + \varepsilon z_j/b, 2\varepsilon)$ ,  $i = 1, \dots, 2^k$ ,  $j = 1, \dots, 2^l$ . Relations (5.10) and (5.11) imply  $t \in G_{i,j}$ . Thus

$$\varepsilon_{k+l}(\mathcal{T}(Q_n)_q, L_p) \leq 2\varepsilon.$$

□

Let  $1 \leq q < p \leq v \leq \infty$ . The following simple inequality

$$\|g\|_p \leq \|g_q\|_q^{1-\mu} \|g\|_v^\mu, \quad \mu := (1/q - 1/p)(1/q - 1/v)^{-1}$$

implies that

$$\varepsilon_k(\mathcal{T}(Q_n)_q, L_p) \leq 2(\varepsilon_k(\mathcal{T}(Q_n)_q, L_v))^\mu. \quad (5.12)$$

**Lemma 5.7.** *Let  $1 < q < p < \infty$  and  $\beta := 1/q - 1/p$ . Then*

$$\varepsilon_k(\mathcal{T}(Q_n)_q, L_p) \ll \begin{cases} (|Q_n|/k)^\beta (\log(4|Q_n|/k))^\beta, & k \leq 2|Q_n|, \\ 2^{-k/(2|Q_n|)}, & k \geq 2|Q_n|. \end{cases}$$

*Proof.* In the case  $1 < q \leq 2 \leq p < \infty$  Lemma 5.7 follows directly from Lemma 5.4. Corollary 2.1 shows that it is sufficient to prove Lemma 5.7 for  $k \leq 2|Q_n|$ . Consider first the case  $2 < q < p < \infty$ . Applying Lemma 5.6 with  $u = 2$ ,  $l = 0$  and using Lemma 5.4 with  $q = 2$  we obtain the required bound. Second, in the case  $1 < q < p < 2$  the required bound follows from inequality (5.12) with  $v = 2$  and Lemma 5.4 with  $p = 2$ . □

Theorem 4.1 and Lemma 5.7 imply the upper bounds in the following theorem.

**Theorem 5.1.** *Let  $1 < q < p < \infty$  and  $a > \beta := 1/q - 1/p$ . Then*

$$\varepsilon_k(\mathbf{W}_q^{a,b}, L_p) \asymp \varepsilon_k(\bar{\mathbf{W}}_q^{a,b}, L_p) \asymp k^{-a} (\log k)^{(d-1)(a+b)}. \quad (5.13)$$

The lower bound in this theorem follows from the lower bound in Theorem 1.7 above.

## 6 Hyperbolic cross polynomials, $d = 2$ , $q = 1$ and $q = \infty$

The construction of the orthonormal basis from the previous section uses classical methods and classical building blocks – an analog of the Dirichlet kernels. That construction works very well for  $L_q$  spaces with  $q \in (1, \infty)$ . However, because of use of the Dirichlet kernels it does not work well in the case  $q = 1$  or  $q = \infty$ . We present here other construction, which is based on the wavelet theory. This construction is taken from [15]. Let  $\delta$  be a fixed number,  $0 < \delta \leq 1/3$ , and let  $\hat{\varphi}(\lambda) = \hat{\varphi}_\delta(\lambda)$ ,  $\lambda \in \mathbb{R}$ , be a sufficiently smooth function (for simplicity, real-valued and even) equal 1 for  $|\lambda| \leq (1 - \delta)/2$ , equal to 0 for  $|\lambda| > (1 + \delta)/2$  and such that the integral translates of its square constitute a partition of unity:

$$\sum_{k \in \mathbb{Z}} (\hat{\varphi}(\lambda + k))^2 = 1, \quad \lambda \in \mathbb{R}. \quad (6.1)$$

It is known that condition (6.1) is equivalent to the following property: The set of functions  $\Phi := \{\varphi(\cdot + l)\}_{l \in \mathbb{Z}}$ , where

$$\varphi(x) = \int_{\mathbb{R}} \hat{\varphi}(\lambda) e^{2\pi i \lambda x} d\lambda,$$

is an orthonormal system on  $\mathbb{R}$ :

$$\int_{\mathbb{R}} \varphi(x + k) \varphi(x + l) dx = \delta_{k,l}, \quad k, l \in \mathbb{Z}. \quad (6.2)$$

Following [15] define

$$\theta(\lambda) := (((\hat{\varphi}(\lambda/2))^2 - (\hat{\varphi}(\lambda))^2)^{1/2}$$

and consider, for  $n = 0, 1, \dots$ , the trigonometric polynomials

$$\Psi_n(x) := 2^{-n/2} \sum_{k \in \mathbb{Z}} \theta(k 2^{-n}) e^{2\pi i k x}. \quad (6.3)$$

Introduce also the following dyadic translates of  $\Psi_n$ :

$$\Psi_{n,j}(x) := \Psi_n(x - (j + 1/2)2^{-n}),$$



and define the sequence of polynomials  $\{T_k\}_{k=0}^\infty$

$$T_0(x) := 1, \quad T_k(x) := \Psi_{n,j}(x) \quad (6.4)$$

if  $k = 2^n + j$ ,  $n = 0, 1, \dots$ ,  $0 \leq j < 2^n$ . Note that  $T_k$  is the trigonometric polynomial such that

$$\hat{T}_k(\nu) = 0 \quad \text{if} \quad |\nu| \geq 2^n(1 + \delta) \quad \text{or} \quad |\nu| \leq 2^{n-1}(1 - \delta). \quad (6.5)$$

It is proved in [15] that the system  $\{T_k\}_{k=0}^\infty$  is a complete orthonormal basis in all  $L_p$ ,  $1 \leq p \leq \infty$  (here,  $L_\infty$  stands for the space of continuous functions) of 1-periodic functions. Also, it is proved in [15] that

$$|\Psi_n(x)| \leq C(\kappa, \delta) 2^{n/2} (2^n |\sin \pi x| + 1)^{-\kappa} \quad (6.6)$$

with  $\kappa$  determined by the smoothness of  $\hat{\varphi}(\lambda)$ . In particular, we can always make  $\kappa > 1$  assuming that  $\hat{\varphi}(\lambda)$  is smooth enough. It is more convenient for us to consider  $2\pi$ -periodic functions. We define  $\mathcal{V} := \{v_k\}_{k=0}^\infty$  with  $v_k(x) := T_k(x/(2\pi))$  for  $x \in [0, 2\pi)$ .

In the multivariate case of  $\mathbf{x} = (x_1, \dots, x_d)$  we define the system  $\mathcal{V}^d$  as the tensor product of the univariate systems  $\mathcal{V}$ . Namely,  $\mathcal{V}^d := \{v_{\mathbf{k}}(\mathbf{x})\}_{\mathbf{k} \in \mathbb{Z}_+^d}$ , where

$$v_{\mathbf{k}}(\mathbf{x}) := \prod_{i=1}^d v_{k_i}(x_i), \quad \mathbf{k} = (k_1, \dots, k_d).$$

Property (6.6) implies the following simple lemma.

**Lemma 6.1.** *We have*

$$\left\| \sum_{\mathbf{k} \in \rho^+(\mathbf{s})} a_{\mathbf{k}} v_{\mathbf{k}} \right\|_\infty \leq C(d, \kappa, \delta) 2^{\|\mathbf{s}\|_1/2} \max_{\mathbf{k}} |a_{\mathbf{k}}|.$$

We use the notation

$$f_{\mathbf{k}} := \langle f, v_{\mathbf{k}} \rangle = (2\pi)^{-d} \int_{\mathbb{T}^d} f(\mathbf{x}) v_{\mathbf{k}}(\mathbf{x}) d\mathbf{x}.$$

Denote

$$Q_n^+ := \{\mathbf{k} = (k_1, \dots, k_d) \in Q_n : k_i \geq 0, i = 1, 2, \dots\}, \quad \theta_n := \{\mathbf{s} : \|\mathbf{s}\|_1 = n\},$$

$$\mathcal{V}(Q_n) := \{f : f = \sum_{\mathbf{k} \in Q_n^+} c_{\mathbf{k}} v_{\mathbf{k}}\}, \quad \mathcal{V}(Q_n)_A := \{f \in \mathcal{V}(Q_n) : \sum_{\mathbf{k} \in Q_n^+} |f_{\mathbf{k}}| \leq 1\}.$$

We prove three inequalities for  $f \in \mathcal{V}(Q_n)$  in the case  $d = 2$ . Theorem 6.1 is a generalized version of the Small Ball Inequality for the system  $\mathcal{V}^2$ .

**Theorem 6.1.** *Let  $d = 2$ . For any  $f \in \mathcal{V}(Q_n)$  we have for  $l \leq n$*

$$\sum_{\mathbf{s} \in \theta_l} \left\| \sum_{\mathbf{k} \in \rho^+(\mathbf{s})} f_{\mathbf{k}} v_{\mathbf{k}} \right\|_1 \leq C(6 + n - l) \|f\|_{\infty},$$

where the constant  $C$  may depend on the choice of  $\hat{\varphi}$ .

**Theorem 6.2.** *Let  $d = 2$ . For any  $f \in \mathcal{V}(Q_n)$  we have*

$$\sum_{\mathbf{k} \in Q_n^+} |f_{\mathbf{k}}| \leq C 2^{n/2} \|f\|_{\infty},$$

where the constant  $C$  may depend on the choice of  $\hat{\varphi}$ .

**Theorem 6.3.** *Let  $d = 2$ . For any  $f \in \mathcal{V}(Q_n)$  we have*

$$\sum_{\mathbf{k} \in Q_n^+} |f_{\mathbf{k}}| \leq C |Q_n|^{1/2} \|f\|_1,$$

where the constant  $C$  may depend on the choice of  $\hat{\varphi}$ .

Proofs of Theorems 6.1 and 6.3 are based on the Riesz products for the hyperbolic cross polynomials (see [19], [20], [26], [28]). Relation (6.5) implies that for  $s_1$  and  $s_2$  greater than 3 the function  $v_{\mathbf{k}}$  with  $\mathbf{k} \in \rho^+(\mathbf{s})$  may have nonzero Fourier coefficients  $\hat{v}_{\mathbf{k}}(\mathbf{m})$  only for

$$\mathbf{m} \in \rho'(\mathbf{s}) := \{\mathbf{m} = (m_1, m_2) : (1 - \delta)2^{s_i-2} < |m_i| < (1 + \delta)2^{s_i-1}, i = 1, 2\}.$$

In other words

$$v_{\mathbf{k}} \in \mathcal{T}(\rho'(\mathbf{s})), \quad \mathbf{k} \in \rho^+(\mathbf{s}).$$

We introduce some more notations. For any two integers  $a \geq 1$  and  $0 \leq b < a$ , we shall denote by  $AP(a, b)$  the arithmetical progression of the form  $al + b$ ,  $l = 0, 1, \dots$ . Set

$$H_n(a, b) := \{\mathbf{s} = (s_1, s_2) : \mathbf{s} \in \mathbb{Z}_+^2, \quad \|\mathbf{s}\|_1 = n, \quad s_1, s_2 \geq a, \quad s_1 \in AP(a, b)\}.$$

For a subspace  $Y$  in  $L_2(\mathbb{T}^d)$  we denote by  $Y^{\perp}$  its orthogonal complement. We need the following lemma on the Riesz product, which is a variant of Lemma 2.1 from [26].

**Lemma 6.2.** *Take any trigonometric polynomials  $t_s \in \mathcal{T}(\rho'(s))$  and form the function*

$$\Phi(\mathbf{x}) := \prod_{s \in H_n(a,b)} (1 + t_s).$$

*Then for any  $a \geq 6$  and any  $0 \leq b < a$  this function admits the representation*

$$\Phi(\mathbf{x}) = 1 + \sum_{s \in H_n(a,b)} t_s(\mathbf{x}) + g(\mathbf{x})$$

*with  $g \in \mathcal{T}(Q_{n+a-6})^\perp$ .*

*Proof.* We prove that for  $\mathbf{k} = (k_1, k_2)$  such that  $|k_1 k_2| \leq 2^{n+a-6}$  we have  $\hat{g}(\mathbf{k}) = 0$ . This proof follows the ideas from [19]. Let  $w(kt)$  denote either  $\cos kt$  or  $\sin kt$ . Then  $g(\mathbf{x})$  contains terms of the form

$$h(\mathbf{x}) = c \prod_{i=1}^m w(k_1^i x_1) w(k_2^i x_2), \quad \mathbf{k}^i \in \rho'(s^i), \quad (6.7)$$

with all  $s^i$ ,  $i = 1, \dots, m$ ,  $m \geq 2$ , distinct. For the sake of simplicity of notations we assume that  $s_1^1 > s_1^2 > \dots > s_1^m$ . Then for  $h(\mathbf{x})$  the frequencies with respect to  $x_1$  have the form

$$k_1 = k_1^1 \pm k_1^2 \pm \dots \pm k_1^m.$$

Therefore, for  $\delta \leq 1/3$  and  $a \geq 6$  we obtain

$$k_1 > (1 - \delta) 2^{s_1^1 - 2} - \sum_{i \geq 1} (1 + \delta) 2^{s_1^1 - 1 - a_i} > 2^{s_1^1 - 3}. \quad (6.8)$$

In the same way for frequencies of the function  $h(\mathbf{x})$  with respect to  $x_2$  we get  $k_2 > 2^{s_2^m - 3}$ . Consequently,

$$k_1 k_2 > 2^{s_1^1 + s_2^m - 6}.$$

In order to complete the proof it remains to observe that for all terms  $h(\mathbf{x})$  of the function  $g(\mathbf{x})$  we have  $m \geq 2$ , which implies  $s_1^1 + s_2^m \geq n + a$ . The lemma is proved. □

**Proof of Theorem 6.1.** For a rectangle  $R \subset \mathbb{Z}_+^2$  denote

$$S_R(f, \mathcal{V}) := \sum_{\mathbf{k} \in R} f_{\mathbf{k}} v_{\mathbf{k}}, \quad \delta_{\mathbf{s}}(f, \mathcal{V}) := S_{\rho^+(\mathbf{s})}(f, \mathcal{V}).$$

It is proved in [15] that  $\|S_R\|_{L_\infty \rightarrow L_\infty} \leq B$ , where  $B$  may depend only on the function  $\hat{\varphi}$ . Define

$$t_{\mathbf{s}} := S_{\rho^+(\mathbf{s})}(\text{sign } \delta_{\mathbf{s}}(f, \mathcal{V})) B^{-1}.$$

Then  $t_{\mathbf{s}} \in \mathcal{T}(\rho'(\mathbf{s}))$  and  $\|t_{\mathbf{s}}\|_\infty \leq 1$ . By Lemma 6.2 with  $n$  replaced by  $l$  and  $a = 6 + n - l$ , where  $n$  is from Theorem 6.1 we obtain

$$\Phi(\mathbf{x}) = 1 + \sum_{\mathbf{s} \in H_l(6+n-l, b)} t_{\mathbf{s}}(\mathbf{x}) + g(\mathbf{x})$$

with  $g \in \mathcal{T}(Q_n)^\perp$ . Clearly,  $\|\Phi\|_1 = 1$ . Therefore, on one hand

$$\langle f, \Phi - 1 \rangle \leq 2\|f\|_\infty. \quad (6.9)$$

On the other hand

$$\begin{aligned} \langle f, \Phi - 1 \rangle &= \sum_{\mathbf{s} \in H_l(6+n-l, b)} \langle f, t_{\mathbf{s}} \rangle = \sum_{\mathbf{s} \in H_l(6+n-l, b)} \langle \delta_{\mathbf{s}}(f, \mathcal{V}), t_{\mathbf{s}} \rangle \\ &= \sum_{\mathbf{s} \in H_l(6+n-l, b)} \langle \delta_{\mathbf{s}}(f, \mathcal{V}), \text{sign } \delta_{\mathbf{s}}(f, \mathcal{V}) \rangle B^{-1} = B^{-1} \sum_{\mathbf{s} \in H_l(6+n-l, b)} \|\delta_{\mathbf{s}}(f, \mathcal{V})\|_1. \end{aligned} \quad (6.10)$$

Thus, for each  $0 \leq b < 6 + n - l$  we have

$$\sum_{\mathbf{s} \in H_l(6+n-l, b)} \|\delta_{\mathbf{s}}(f, \mathcal{V})\|_1 \leq 2B\|f\|_\infty \quad (6.11)$$

This easily implies the conclusion of Theorem 6.1.

**Proof of Theorem 6.2.** Theorem 6.2 is a corollary of Theorem 6.1. Indeed, by Lemma 6.1 we get

$$\sum_{\mathbf{k} \in \rho^+(\mathbf{s})} |f_{\mathbf{k}}| = \langle \delta_{\mathbf{s}}(f, \mathcal{V}), \sum_{\mathbf{k} \in \rho^+(\mathbf{s})} (\text{sign } f_{\mathbf{k}}) v_{\mathbf{k}} \rangle \leq C \|\delta_{\mathbf{s}}(f, \mathcal{V})\|_1 2^{\|\mathbf{s}\|_1/2}.$$

Thus, by Theorem 6.1 we get

$$\sum_{\mathbf{k} \in Q_n^+} |f_{\mathbf{k}}| \leq C \sum_{l \leq n} (6 + n - l) 2^{l/2} \|f\|_\infty \ll 2^{n/2} \|f\|_\infty,$$

which completes the proof of Theorem 6.2.

**Proof of Theorem 6.3.** We begin with some auxiliary results. The following simple remark is from [28].

**Remark 6.1.** *For any real numbers  $y_l$  such that  $|y_l| \leq 1$ ,  $l = 1, \dots, N$  we have ( $i^2 = -1$ )*

$$\left| \prod_{l=1}^N \left( 1 + \frac{iy_l}{\sqrt{N}} \right) \right| \leq C.$$

We now prove two lemmas, which are analogs of Lemmas 2.2 and 2.3 from [28]. Denote

$$E_{Q_n}^\perp(f)_p := \inf_{g \in \mathcal{T}(Q_n)^\perp} \|f - g\|_p.$$

**Lemma 6.3.** *For any function  $f$  of the form*

$$f = \sum_{\mathbf{s} \in H_n(a,b)} t_{\mathbf{s}}$$

*with  $a \geq 6$ ,  $0 \leq b < a$ , where  $t_{\mathbf{s}}$ ,  $\mathbf{s} \in H_n(a,b)$ , is a real trigonometric polynomial in  $\mathcal{T}(\rho'(\mathbf{s}))$  such that  $\|t_{\mathbf{s}}\|_\infty \leq 1$ , we have*

$$E_{Q_{n+a-6}}^\perp(f)_\infty \leq C(1 + n/a)^{1/2}$$

*with  $C$  depending only on  $\hat{\varphi}$ .*

*Proof.* Let us form the function

$$RP(f) := \text{Im} \prod_{\mathbf{s} \in H_n(a,b)} (1 + it_{\mathbf{s}}(1 + n/a)^{-1/2}),$$

which is an analog of the Riesz product. Then by Remark 6.1 we have

$$\|RP(f)\|_\infty \leq C. \tag{6.12}$$

Lemma 6.2 provides the representation

$$RP(f) = (1 + n/a)^{-1/2} \sum_{\mathbf{s} \in H_n(a,b)} t_{\mathbf{s}} + g, \quad g \in \mathcal{T}(Q_{n+a-6}). \tag{6.13}$$

Combining (6.12) and (6.13) we obtain the statement of Lemma 6.3.  $\square$

**Remark 6.2.** *It is clear that in Lemma 6.3 we can drop the assumption that the  $t_{\mathbf{s}}$  are real polynomials.*

**Lemma 6.4.** *For any function  $f$  of the form*

$$f = \sum_{\mathbf{s} \in \theta_n} t_{\mathbf{s}}, \quad t_{\mathbf{s}} \in \mathcal{T}(\rho'(\mathbf{s})), \quad \|t_{\mathbf{s}}\|_{\infty} \leq 1,$$

*we have for any  $a \geq 6$*

$$E_{Q_{n+a-6}}^{\perp}(f)_{\infty} \leq Ca(1 + n/a)^{1/2}$$

*Proof.* Let us introduce some more notations. Denote

$$\theta_{n,a} := \{\mathbf{s} \in \theta_n : \text{ either } s_1 < a \text{ or } s_2 < a\}.$$

Then

$$f = \sum_{\mathbf{s} \in \theta_n} t_{\mathbf{s}} = \sum_{\mathbf{s} \in \theta_{n,a}} t_{\mathbf{s}} + \sum_{b=0}^{a-1} \sum_{\mathbf{s} \in H_n(a,b)} t_{\mathbf{s}}$$

and

$$E_{Q_{n+a-6}}^{\perp}(f)_{\infty} \leq \sum_{\mathbf{s} \in \theta_{n,a}} \|t_{\mathbf{s}}\|_{\infty} + \sum_{b=0}^{a-1} E_{Q_{n+a-6}}^{\perp} \left( \sum_{\mathbf{s} \in H_n(a,b)} t_{\mathbf{s}} \right)_{\infty}.$$

Using the assumption  $\|t_{\mathbf{s}}\|_{\infty} \leq 1$ , Lemma 6.3, and Remark 6.2 we get from here the required estimate.  $\square$

We now proceed to the proof of Theorem 6.3. For  $l \in [0, n]$  consider

$$t_{\mathbf{s}}^1 := \sum_{\mathbf{k} \in \rho^+(\mathbf{s})} (\text{sign } f_{\mathbf{k}}) v_{\mathbf{k}}, \quad M_l := \max_{\mathbf{s} \in \theta_l} \|t_{\mathbf{s}}^1\|_{\infty}, \quad t_{\mathbf{s}} := t_{\mathbf{s}}^1 / M_l.$$

By Lemma 6.1

$$M_l \ll 2^{l/2}.$$

Applying Lemma 6.4 with  $a = 6 + n - l$  we get

$$\begin{aligned} \sum_{\mathbf{s} \in \theta_l} \sum_{\mathbf{k} \in \rho^+(\mathbf{s})} |f_{\mathbf{k}}| &= \langle f, \sum_{\mathbf{s} \in \theta_l} t_{\mathbf{s}}^1 \rangle = M_l \langle f, \sum_{\mathbf{s} \in \theta_l} t_{\mathbf{s}} \rangle \\ &\ll 2^{l/2} E_{Q_n}^{\perp} \left( \sum_{\mathbf{s} \in \theta_l} t_{\mathbf{s}} \right)_{\infty} \|f\|_1 \ll 2^{l/2} (6 + n - l) \left( 1 + \frac{n}{6 + n - l} \right)^{1/2}. \end{aligned}$$

Summing up over  $l \leq n$ , we complete the proof of Theorem 6.3.

**Lemma 6.5.** *Let  $2 \leq p < \infty$ . Let  $\mathcal{V}_n^1 := \{v_{\mathbf{k}} : \mathbf{k} \in Q_n\}$ . Then*

$$\sigma_m(\mathcal{V}(Q_n)_A, \mathcal{V}_n^1)_p \ll |Q_n|^{1/2-1/p} m^{1/p-1}. \quad (6.14)$$

*Proof.* Note, that for  $f \in \mathcal{V}(Q_n)_A$  we easily obtain that  $\|f\|_2 \leq 1$  and  $\|f\|_\infty \ll 2^{n/2}$ , which, in turn, implies for  $2 \leq p \leq \infty$

$$\|f\|_p \ll 2^{n(1/2-1/p)}. \quad (6.15)$$

Thus, it is sufficient to prove (6.14) for big enough  $m$ .

First, we prove the lemma for  $\Delta Q_n$  instead of  $Q_n$ . Then  $f \in \mathcal{V}(\Delta Q_n)_A$  has a representation

$$f = \sum_{\mathbf{k} \in \Delta Q_n} f_{\mathbf{k}} v_{\mathbf{k}}, \quad \sum_{\mathbf{k} \in \Delta Q_n} |f_{\mathbf{k}}| \leq 1.$$

Using the fact that the system  $\mathcal{V}^d$  is orthonormal we obtain by Lemma 5.2 with  $m_1 := [m/2]$  that

$$\sigma_{m_1}(f, \mathcal{V}^d)_2 \leq (m_1 + 1)^{-1/2}. \quad (6.16)$$

For a set  $\Lambda$  denote

$$\mathcal{V}(\Lambda)_q := \{f : f = \sum_{\mathbf{k} \in \Lambda \cap \mathbb{Z}_+^d} f_{\mathbf{k}} v_{\mathbf{k}}, \quad \|f\|_q \leq 1\}.$$

Next, we estimate the best  $m_1$ -term approximation of  $g \in \mathcal{V}(\Delta Q_n)_2$  in  $L_p$ ,  $2 < p < \infty$ . We apply Lemma 5.2 to each set of  $g_{\mathbf{k}}$ ,  $\mathbf{k} \in \rho^+(\mathbf{s})$ ,  $\|\mathbf{s}\|_1 = n$  with  $m_{\mathbf{s}} := [m_1/m_n]$ ,  $m_n := [\Delta Q_n 2^{-n}]$ , assuming that  $n \geq C$  with the absolute constant  $C$  large enough to guarantee  $m_n \geq 1$ . Denote  $G_{\mathbf{s}}$  the set of cardinality  $|G_{\mathbf{s}}| = m_{\mathbf{s}}$  of  $\mathbf{k}$  from  $\rho^+(\mathbf{s})$  with largest  $|g_{\mathbf{k}}|$ . Then by Lemmas 5.2 and 6.1 we obtain

$$\left\| \sum_{\mathbf{k} \in \rho^+(\mathbf{s}) \setminus G_{\mathbf{s}}} g_{\mathbf{k}} v_{\mathbf{k}} \right\|_\infty \ll 2^{n/2} (m_{\mathbf{s}} + 1)^{-1/2} \|\delta_{\mathbf{s}}(g, \mathcal{V})\|_2.$$

Applying simple inequality for  $2 \leq p \leq \infty$

$$\|f\|_p \leq \|f\|_2^\alpha \|f\|_\infty^{1-\alpha}, \quad \alpha = 2/p$$

we obtain

$$\left\| \sum_{\mathbf{k} \in \rho(\mathbf{s}) \setminus G_{\mathbf{s}}} g_{\mathbf{k}} v_{\mathbf{k}} \right\|_p \ll (|\Delta Q_n|/m_1)^{1/2-1/p} \|\delta_{\mathbf{s}}(g, \mathcal{V})\|_2. \quad (6.17)$$

Inequality (5.2) implies easily a similar inequality for  $\mathcal{V}^d$  for  $2 \leq p < \infty$

$$\|f\|_p \ll \left( \sum_{\mathbf{s}} \|\delta_{\mathbf{s}}(f, \mathcal{V})\|_p^2 \right)^{1/2}. \quad (6.18)$$

Combining (6.16), (6.17), and (6.18) we complete the proof of Lemma 6.5 in the case of  $\Delta Q_n$ .

We now derive the general case of  $Q_n$  from the above considered case of  $\Delta Q_l$ . Set

$$\mu := \frac{1}{2} \left( \frac{1}{2} - \frac{1}{p} \right) \left( 1 - \frac{1}{p} \right)^{-1}$$

and denote by  $l_0$  the smallest  $l$  satisfying

$$m2^{-\mu(n-l)} \geq 1. \quad (6.19)$$

Then for  $f \in \mathcal{V}(Q_n)_A$  we have by (6.15) and (6.19)

$$f_0 := \sum_{\mathbf{k} \in Q_{l_0}^+} f_{\mathbf{k}} v_{\mathbf{k}}, \quad \|f_0\|_p \ll 2^{l_0(1/2-1/p)} \ll 2^{n(1/2-1/p)} m^{2(1/p-1)}. \quad (6.20)$$

For  $l > l_0$  define  $m_l := \lceil m2^{-\mu(n-l)} \rceil \geq 1$ . Then

$$m' := \sum_{l_0 < l \leq n} m_l \leq C(p)m,$$

and

$$\sigma_{m'}(f - f_0, \mathcal{V}_n^1)_p \ll \sum_{l_0 < l \leq n} |\Delta Q_l|^{1/2-1/p} m_l^{1/p-1} \ll |Q_n|^{1/2-1/p} m^{1/p-1}. \quad (6.21)$$

Combining (6.20) – (6.21) we complete the proof of Lemma 6.5. □

Lemma 6.5 and Theorem 6.3 imply.

**Lemma 6.6.** *Let  $2 \leq p < \infty$ . Let  $\mathcal{V}_n^1 := \{v_{\mathbf{k}} : \mathbf{k} \in Q_n\}$ . Then*

$$\sigma_m(\mathcal{V}(Q_n)_1, \mathcal{V}_n^1)_p \ll (|Q_n|/m)^{1-1/p}.$$

Lemma 6.5 and Theorem 6.2 imply.



**Lemma 6.7.** *Let  $2 \leq p < \infty$ . Let  $\mathcal{V}_n^1 := \{v_{\mathbf{k}} : \mathbf{k} \in Q_n\}$ . Then*

$$\sigma_m(\mathcal{V}(Q_n)_\infty, \mathcal{V}_n^1)_p \ll n^{-1/2}(|Q_n|/m)^{1-1/p}.$$

We now apply the second step of the strategy described in Section 4. Theorem 4.2, Remark 4.2 and Lemma 6.6 imply the following lemma.

**Lemma 6.8.** *Let  $2 \leq p < \infty$  and  $\beta := 1 - 1/p$ . Then*

$$\varepsilon_k(\mathcal{V}(Q_n)_1, L_p) \ll \begin{cases} (|Q_n|/k)^\beta (\log(4|Q_n|/k))^\beta, & k \leq 2|Q_n|, \\ 2^{-k/(2|Q_n|)}, & k \geq 2|Q_n|. \end{cases}$$

Theorem 4.2, Remark 4.2 and Lemma 6.7 imply the following lemma.

**Lemma 6.9.** *Let  $2 \leq p < \infty$  and  $\beta := 1 - 1/p$ . Then*

$$\varepsilon_k(\mathcal{V}(Q_n)_\infty, L_p) \ll \begin{cases} n^{-1/2}(|Q_n|/k)^\beta (\log(4|Q_n|/k))^\beta, & k \leq 2|Q_n|, \\ n^{-1/2}2^{-k/(2|Q_n|)}, & k \geq 2|Q_n|. \end{cases}$$

The fact that  $\mathcal{T}(Q_n) \subset \mathcal{V}(Q_{n+6})$ , Corollary 2.1 and inequality (5.12) allow us to derive the following results from Lemmas 6.8 and 6.9.

**Theorem 6.4.** *Let  $1 < p < \infty$  and  $\beta := 1 - 1/p$ . Then*

$$\varepsilon_k(\mathcal{T}(Q_n)_1, L_p) \ll \begin{cases} (|Q_n|/k)^\beta (\log(4|Q_n|/k))^\beta, & k \leq 2|Q_n|, \\ 2^{-k/(2|Q_n|)}, & k \geq 2|Q_n|. \end{cases}$$

**Theorem 6.5.** *Let  $2 \leq p < \infty$  and  $\beta := 1 - 1/p$ . Then*

$$\varepsilon_k(\mathcal{T}(Q_n)_\infty, L_p) \ll \begin{cases} n^{-1/2}(|Q_n|/k)^\beta (\log(4|Q_n|/k))^\beta, & k \leq 2|Q_n|, \\ n^{-1/2}2^{-k/(2|Q_n|)}, & k \geq 2|Q_n|. \end{cases}$$

## 7 Hyperbolic cross polynomials, $p = \infty$

We now discuss a more difficult and more interesting case  $p = \infty$ . Denote

$$\|f\|_A := \sum_{\mathbf{k}} |\hat{f}(\mathbf{k})|, \quad \hat{f}(\mathbf{k}) := (2\pi)^{-d} \int_{\mathbb{T}^d} f(\mathbf{x}) e^{-i(\mathbf{k}, \mathbf{x})} d\mathbf{x}.$$

The following Theorem 7.1 is from [32] (see Theorem 2.6). We use it in this paper. Let as above

$$\Pi(\mathbf{N}, d) := \{(a_1, \dots, a_d) \in \mathbb{Z}^d : |a_j| \leq N_j, j = 1, \dots, d\},$$

where  $N_j$  are nonnegative integers and  $\mathbf{N} := (N_1, \dots, N_d)$ . We denote

$$\mathcal{T}(\mathbf{N}, d) := \mathcal{T}(\Pi(\mathbf{N}, d)) = \{t : t = \sum_{\mathbf{k} \in \Pi(\mathbf{N}, d)} c_{\mathbf{k}} e^{i(\mathbf{k}, \mathbf{x})}\}.$$

Then

$$\dim \mathcal{T}(\mathbf{N}, d) = \prod_{j=1}^d (2N_j + 1) =: \vartheta(\mathbf{N}).$$

For a nonnegative integer  $m$  denote  $\bar{m} := \max(m, 1)$ .

**Theorem 7.1.** *There exist constructive greedy-type approximation methods  $G_m^p(\cdot)$ , which provide  $m$ -term polynomials with respect to  $\mathcal{T}^d$  with the following properties: for  $2 \leq p < \infty$*

$$\|f - G_m^p(f)\|_p \leq C_1(d)(\bar{m})^{-1/2} p^{1/2} \|f\|_A,$$

and for  $p = \infty$ ,  $f \in \mathcal{T}(\mathbf{N}, d)$

$$\|f - G_m^\infty(f)\|_\infty \leq C_2(d)(\bar{m})^{-1/2} (\ln \vartheta(\mathbf{N}))^{1/2} \|f\|_A.$$

We will use a version of Theorem 7.1, which follows from the proof of Theorem 7.1 in [32] and a simple inequality

$$\|f\|_A \leq \|f\|_2 (\#\{\mathbf{k} : \hat{f}(\mathbf{k}) \neq 0\})^{1/2}.$$

**Theorem 7.2.** *Let  $\Lambda \subset \Pi(\mathbf{N}, d)$  with  $N_j = 2^n$ ,  $j = 1, \dots, d$ . There exist constructive greedy-type approximation methods  $G_m^\infty(\cdot)$ , which provide  $m$ -term polynomials with respect to  $\mathcal{T}^d$  with the following properties: for  $f \in \mathcal{T}(\Lambda)$  we have  $G_m^\infty(f) \in \mathcal{T}(\Lambda)$  and*

$$\|f - G_m^\infty(f)\|_\infty \leq C_3(d)(\bar{m})^{-1/2} n^{1/2} |\Lambda|^{1/2} \|f\|_2.$$

We now prove the following lemma. Let  $\mathcal{D}_n^2 := \{u_{\mathbf{k}} : \mathbf{k} \in Q_n\} \cup \{e^{i(\mathbf{k}, \mathbf{x})} : \mathbf{k} \in Q_n\}$ .

**Lemma 7.1.** *Let  $1 < q \leq 2$ . Then*

$$\sigma_m(\mathcal{T}(Q_n)_q, \mathcal{D}_n^2)_\infty \ll n^{1/2} (|Q_n|/m)^{1/q}.$$

*Proof.* Take  $f \in \mathcal{T}(Q_n)$ . Applying first Lemma 5.3 with  $p = 2$  and  $[m/2]$  and, then, applying Theorem 7.2 with  $\Lambda = Q_n$  and  $[m/2]$  we obtain

$$\sigma_m(f, \mathcal{D}_n^2)_\infty \ll n^{1/2}(|Q_n|/m)^{1/q} \|f\|_q,$$

which proves the lemma.  $\square$

Theorem 4.2, Remark 4.2 and Lemma 7.1 imply the following lemma.

**Theorem 7.3.** *Let  $1 < q \leq 2$ . Then*

$$\varepsilon_k(\mathcal{T}(Q_n)_q, L_\infty) \ll \begin{cases} n^{1/2}(|Q_n|/k)^{1/q} (\log(4|Q_n|/k))^{1/q}, & k \leq 2|Q_n|, \\ n^{1/2} 2^{-k/(2|Q_n|)}, & k \geq 2|Q_n|. \end{cases}$$

Let us discuss the case  $q = 1$ , which is not covered by Theorem 7.3. In this case we restrict ourselves to  $d = 2$ . It is easy to see that  $\mathcal{V}(Q_n) \subset \mathcal{T}(Q_n)$ . Let  $\mathcal{V}_n^2 := \{v_{\mathbf{k}} : \mathbf{k} \in Q_n\} \cup \{e^{i(\mathbf{k}, \mathbf{x})} : \mathbf{k} \in Q_n\}$ . Then we have the following analog of Lemma 7.1.

**Lemma 7.2.** *We have for  $d = 2$*

$$\sigma_m(\mathcal{V}(Q_n)_1, \mathcal{V}_n^2)_\infty \ll n^{1/2}|Q_n|/m.$$

*Proof.* Take  $f \in \mathcal{V}(Q_n)$ . Applying first Lemma 6.6 with  $p = 2$  and  $[m/2]$  and, then, applying Theorem 7.2 with  $\Lambda = Q_n$  and  $[m/2]$  we obtain

$$\sigma_m(f, \mathcal{V}_n^2)_\infty \ll n^{1/2}(|Q_n|/m) \|f\|_q,$$

which proves the lemma.  $\square$

In the same way as above we derive the following result on the entropy numbers from Lemma 7.2.

**Theorem 7.4.** *We have for  $d = 2$*

$$\varepsilon_k(\mathcal{T}(Q_n)_1, L_\infty) \ll \begin{cases} n^{1/2}(|Q_n|/k) \log(4|Q_n|/k), & k \leq 2|Q_n|, \\ n^{1/2} 2^{-k/(2|Q_n|)}, & k \geq 2|Q_n|. \end{cases}$$

## 8 The upper bounds for function classes

In this section we provide a proof of the upper bounds in Theorems 1.3 – 1.8, except Theorem 1.5, formulated in the Introduction. The proof uses Theorem 4.1, which is a general result on the  $\bar{\mathbf{W}}_X^{a,b}\{X_n\}$ . We apply this theorem in the case  $X = L_q$ ,  $1 \leq q \leq \infty$ ,  $X_n = \mathcal{T}(Q_n)$ . An important ingredient of Theorem 4.1 is the assumption that subspaces  $\{X_n\}$  satisfy (4.1). The main work of this paper is devoted to establishing (4.1) in the case  $X_n = \mathcal{T}(Q_n)$ ,  $X = L_q$ ,  $Y = L_p$ , for different parameters  $1 \leq q, p \leq \infty$ . We now indicate which results are used to obtain the appropriate versions of the (4.1) needed for the proof of the upper bounds in the corresponding theorem.

**Proof of Theorem 1.3.** The case  $p = 1$  follows from Remark 4.1. The case  $1 < p < \infty$  follows from Theorem 6.4, which provides (4.1) with  $\alpha = 0$ ,  $\beta = \gamma = 1 - 1/p$ .

**Proof of Theorem 1.4.** It follows from Theorem 7.4, which provides (4.1) with  $\alpha = 1/2$ ,  $\beta = \gamma = 1$ .

**Proof of Theorem 1.6.** It follows from Theorem 6.5, which provides (4.1) with  $\alpha = -1/2$ ,  $\beta = \gamma = 1 - 1/p$  for  $p \in [2, \infty)$ .

**Proof of Theorem 1.7.** It follows from Remark 4.1.

**Proof of Theorem 1.8.** It follows from Theorem 7.3, which provides (4.1) with  $\alpha = 1/2$ ,  $\beta = \gamma = 1/q$ .

## 9 Some properties of classes $\mathbf{W}_q^{a,b}$ and $\bar{\mathbf{W}}_q^{a,b}$

In this section we give some embedding type properties of classes  $\mathbf{W}_q^{a,b}$  and  $\bar{\mathbf{W}}_q^{a,b}$ . As we already mentioned above we have the following two properties.

**Proposition 9.1.** *For  $1 < q < \infty$  the classes  $\mathbf{W}_q^{a,b}$  and  $\bar{\mathbf{W}}_q^{a,b}$  are equivalent.*

**Proposition 9.2.** *For  $1 < q < \infty$  we have  $\mathbf{W}_{q,\alpha}^r \hookrightarrow \mathbf{W}_q^{r,0}$ .*

The Nikol'skii type inequalities from Theorem 2.3 imply the following embeddings.

**Proposition 9.3.** *For  $1 \leq q < p < \infty$  and  $a > \beta := 1/q - 1/p$  we have  $\mathbf{W}_q^{a,b} \hookrightarrow \mathbf{W}_p^{a-\beta,b}$  and  $\bar{\mathbf{W}}_q^{a,b} \hookrightarrow \bar{\mathbf{W}}_p^{a-\beta,b}$ .*

It is well known (see [20] and [25]) that for classes  $\mathbf{W}_{q,\alpha}^r$  and other classes with mixed smoothness approximation by the hyperbolic cross trigonometric polynomials  $\mathcal{T}(Q_n)$  plays the same role as approximation by classical

trigonometric polynomials plays in approximation of the univariate smoothness classes. There are two classical ways to describe compact sets, which are of interest in approximation theory and numerical analysis (see, for instance, [3]). The first way to describe such classes uses the notion of smoothness. A typical example of such classes are classes  $\mathbf{W}_{q,\alpha}^r$ . The second way to describe such classes is through approximation. In our context these would be the classes

$$\mathbf{A}_q^{a,b} := \{f \in L_q : E_{Q_n}(f)_q \leq 2^{-an} n^{(d-1)b}, \quad n = 1, 2, \dots, \quad \|f\|_q \leq 1\},$$

where

$$E_{Q_n}(f)_q := \inf_{t \in \mathcal{T}(Q_n)} \|f - t\|_q$$

is the best approximation of  $f$  by the hyperbolic cross polynomials from  $\mathcal{T}(Q_n)$  in the  $L_q$  norm.

**Proposition 9.4.** *For  $1 \leq q \leq \infty$  classes  $\bar{\mathbf{W}}_q^{a,b}$  and  $\mathbf{A}_q^{a,b}$  are equivalent.*

*Proof.* Let  $f \in \bar{\mathbf{W}}_q^{a,b}$ . Then by the definition of this class  $f$  has a representation

$$f = \sum_{j=1}^{\infty} t_j, \quad t_j \in \mathcal{T}(Q_j), \quad \|t_j\|_q \leq 2^{-aj} j^{b(d-1)}.$$

Therefore,

$$E_{Q_n}(f)_q \leq \left\| \sum_{j>n} t_j \right\|_q \ll 2^{-an} n^{b(d-1)},$$

which implies  $\bar{\mathbf{W}}_q^{a,b} \hookrightarrow \mathbf{A}_q^{a,b}$ . Let now  $f \in \mathbf{A}_q^{a,b}$  and let  $u_n \in \mathcal{T}(Q_n)$  be such that

$$\|f - u_n\|_q = E_{Q_n}(f)_q.$$

Define

$$t_1 := u_1, \quad t_n := u_n - u_{n-1}, \quad n = 2, 3, \dots$$

Then

$$f = \sum_{n=1}^{\infty} t_n, \quad t_n \in \mathcal{T}(Q_n),$$

with

$$\|t_1\|_q \leq 2, \quad \|t_n\|_q \leq 3(2^{-an} n^{b(d-1)}), \quad n = 2, 3, \dots,$$

which implies  $\mathbf{A}_q^{a,b} \hookrightarrow \bar{\mathbf{W}}_q^{a,b}$ . □

Here is a nontrivial embedding type inequality, which is a direct corollary of Lemma 3.3 of Chapter 1 from [20].

**Proposition 9.5.** *For  $1 \leq q < p < \infty$  we have*

$$E_{Q_n}(f)_p \leq C(q, p, d) \left( \sum_{j>n} 2^{pj\beta} E_{Q_n}(f)_q^p \right)^{1/p}.$$

We formulate one more nontrivial inequality (see Lemma 4.3 in [33]), which might be useful in the further study of classes  $\mathbf{W}_1^{a,b}$  and  $\bar{\mathbf{W}}_1^{a,b}$ .

**Lemma 9.1.** *Let  $1 < p < \infty$ . For any  $\epsilon > 0$  there exists a constant  $C(\epsilon, d, p)$  such that for each  $t \in \mathcal{T}(Q_n)$  we have*

$$\sum_{\|s\|_1 \leq n} \|\delta_s(t)\|_p \leq C(\epsilon, d, p) n^\epsilon 2^{\beta n} \|t\|_1, \quad \beta = 1 - 1/p.$$

## References

- [1] R.F. Bass, Probability estimates for multiparameter Brownian processes, *Ann. Probab.*, **16** (1988), 251–264.
- [2] E.S. Belinsky, Estimates of entropy numbers and Gaussian measures for classes of functions with bounded mixed derivative, *J. Approx. Theory*, **93** (1998), 114–127.
- [3] R.A. DeVore, Nonlinear approximation *Acta Numerica* **7** (1998), 51–150.
- [4] Ding Dung, Approximation of multivariate functions by means of harmonic analysis, *Hab. Dissertation* (1985), Moscow, MGU.
- [5] Ding Dung, V.N. Temlyakov, and T. Ullrich, Hyperbolic Cross Approximation, arXiv:1601.03978v1 [math.NA] 15 Jan 2016.
- [6] E.D. Gluskin, Extremal properties of orthogonal parallelepipeds and their applications to the geometry of Banach spaces, *Mat. Sb.*, **136** (1988), 85–96.
- [7] K. Höllig Diameters of classes of smooth functions in Quantitative Approximation, Academic Press, New York, 1980 163–176

- [8] B.S. Kashin, On certain properties of the space of trigonometric polynomials with the uniform norm, *Trudy Mat. Inst. Steklov*, **145** (1980), 111–116; English transl. in *Proc. Steklov Inst. Math.*, **145** (1981).
- [9] B. S. Kashin and V. N. Temlyakov, On best  $m$ -terms approximations and the entropy of sets in the space  $L^1$ , *Mat. Zametki*, **56** (1994), 57–86; English transl. in *Math. Notes*, **56** (1994), 1137–1157.
- [10] B.S. Kashin and V.N. Temlyakov, Estimate of approximate characteristics for classes of functions with bounded mixed derivative, *Math. Notes*, **58** (1995), 1340–1342.
- [11] B.S. Kashin and V.N. Temlyakov, The volume estimates and their applications, *East J. Approx.*, **9** (2003), 469–485.
- [12] J. Kuelbs and W.V. Li, Metric entropy and the small ball problem for Gaussian measures, *J. Functional Analysis*, **116** (1993), 133–157.
- [13] M.A. Lifshits and B.S. Tsirelson, Small deviations of Gaussian fields, *Teor. Probab. Appl.*, **31** (1986), 557–558.
- [14] V.E. Maiorov On various widths of the class  $H_p^r$  in the space  $L_q$  *Izv. Akad. Nauk SSSR Ser. Mat.* **42** (1978), 773–788; English transl. in *Math. USSR-Izv.* **13** (1979).
- [15] D. Offin and K. Oskolkov, A note on orthonormal polynomial bases and wavelets, *Constructive Approx.* **9** (1993), 319–325.
- [16] C. Schütt, Entropy numbers of diagonal operators between symmetric Banach spaces, *J. Approx. Theory*, **40** (1984), 121–128.
- [17] S.A. Smolyak, The  $\epsilon$ -entropy of the classes  $E_s^{\alpha k}(B)$  and  $W_s^\alpha(B)$  in the metric  $L_2$ , *Dokl. Akad. Nauk SSSR*, **131** (1960), 30–33.
- [18] M. Talagrand, The small ball problem for the Brownian sheet, *Ann. Probab.*, **22** (1994), 1331–1354.
- [19] V.N Temlyakov, Approximation of periodic functions of several variables with bounded mixed difference, *Mat. Sb.*, **133** (1980), 65–85; English transl. in *Math. USSR Sbornik* **41** (1982).

- [20] V.N. Temlyakov, Approximation of functions with bounded mixed derivative, *Trudy MIAN*, **178** (1986), 1–112. English transl. in *Proc. Steklov Inst. Math.*, **1** (1989).
- [21] V.N. Temlyakov, On estimates of  $\epsilon$ -entropy and widths of classes of functions with bounded mixed derivative or difference, *Dokl. Akad. Nauk SSSR*, **301** (1988), 288–291; English transl. in *Soviet Math. Dokl.*, **38**, 84–87.
- [22] V.N. Temlyakov, Estimates of the asymptotic characteristics of classes of functions with bounded mixed derivative or difference, *Trudy Matem. Inst. Steklov*, **189** (1989), 138–168; English transl. in *Proceedings of the Steklov Institute of Mathematics*, 1990, Issue 4, 161–197.
- [23] V.N. Temlyakov, Bilinear Approximation and Related Questions, *Trudy Mat. Inst. Steklov*, **194** (1992), 229–248; English transl. in *Proc. Steklov Inst. of Math.*, **4** (1993), 245–265.
- [24] V.N. Temlyakov, Estimates of Best Bilinear Approximations of Functions and Approximation Numbers of Integral Operators, *Mat. Zametki*, **51** (1992), 125–134; English transl. in *Math. Notes*, **51** (1992), 510–517.
- [25] V.N. Temlyakov, *Approximation of periodic functions*, Nova Science Publishes, Inc., New York., 1993.
- [26] V.N. Temlyakov, An inequality for trigonometric polynomials and its application for estimating the entropy numbers, *J. Complexity*, **11** (1995), 293–307.
- [27] V.N. Temlyakov, Some inequalities for multivariate Haar polynomials, *East J. Approx.*, **1** (1995), 61–72.
- [28] V.N. Temlyakov, On two problems in the multivariate approximation, *East J. Approx.*, **4** (1998), 505–514.
- [29] V.N. Temlyakov, Greedy Algorithms with Regards to Multivariate Systems with Special Structure, *Constructive Approximation*, **16** (2000), 399–425.



- [30] V.N. Temlyakov, Greedy approximation, Cambridge University Press, 2011.
- [31] V.N. Temlyakov, An inequality for the entropy numbers and its application, J. Approximation Theory, **173** (2013), 110–121.
- [32] V.N. Temlyakov, Constructive sparse trigonometric approximation and other problems for functions with mixed smoothness, Matem. Sb. **206** (2015), 131–160; arXiv: 1412.8647v1 [math.NA] 24 Dec 2014, 1–37.
- [33] V.N. Temlyakov, Constructive sparse trigonometric approximation for functions with small mixed smoothness, arXiv:1503.0282v1 [math.NA] 1 Mar 2015.
- [34] H. Triebel, Bases in Function Spaces, Sampling, Discrepancy, Numerical Integration, European Mathematical Society, Germany, 2010.